

# INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

## IN MEDIZIN UND BIOLOGIE

Offizielles Organ  
der Deutschen Gesellschaft für  
Medizinische Informatik,  
Biometrie und Epidemiologie (GMDS) e.V.

Persönliches Exemplar für Mitglieder  
der Deutschen Gesellschaft für  
Medizinische Informatik,  
Biometrie und Epidemiologie (GMDS) e.V.,  
darf nicht in öffentlichen Bibliotheken  
eingestellt werden.

Urban & Fischer Verlag  
Verlag Eugen Ulmer Stuttgart

**Band 30**  
**Heft 2/1999**  
ISSN 0943-5581

# INFORMATIK BIOMETRIE und EPIDEMIOLOGIE

## IN MEDIZIN UND BIOLOGIE

Urban & Fischer Verlag Jena  
Verlag Eugen Ulmer Stuttgart

Schriftleitung:  
Prof. Dr. Markus Löffler, Leipzig

### Herausgeber

P. Bauer (Wien) · M. Blettner (Heidelberg) · J. Dudeck (Gießen) · U. Feldmann (Homburg) · H. Geidel (Stuttgart)  
R. Haux (Heidelberg) · W. Lehmacher (Köln) · M. Löffler (Leipzig) · J. Michaelis (Mainz) · H. Thöni (Hohenheim)  
J. Vollmar (Mannheim) · H.-E. Wichmann (München)

### Wissenschaftlicher Beirat

H. Becher (Heidelberg) · J. Berger (Hamburg) · W. van Eimeren (Neuherberg) · U. Ferner (Basel)  
H. Haußmann (Hohenheim) · H.-W. Hense (Münster) · P. Jensch (Oldenburg) · K.-H. Jöckel (Essen)  
C. O. Köhler (Heidelberg) · W. Köhler (Gießen) · W. Maurer (Basel) · R. Mösges (Aachen)  
O. Richter (Braunschweig) · H. Rundfeldt (Hannover) · M. Schumacher (Freiburg) · S. Stiehl (Hamburg)  
Th. Tolxdorff (Berlin) · H.-D. Unkelbach (Geisenheim) · H. F. Utz (Hohenheim) · J. Wahrendorf (Heidelberg)

## Inhaltsverzeichnis 2/1999

Modellbasierte Prädiktion der Diabetesprävalenz	P. Heinke, E. Salzsieder	49
The Reliability Evaluation of Nuclear Imaging Data: An overview on classical ANOVA based approaches	F. Krummenauer, G. Förster	60
Effekte fehlerhafter Expositions- und Confounder- bestimmungen in epidemiologischen Studien	G. Rippin	72
Buchbesprechungen/Bookreviews		83

### Redaktion

**Schriftleiter** (verantwortlich im Sinne des Presserechts): Prof. Dr. M. Löffler  
Universität Leipzig, Institut für Medizinische Informatik, Statistik und Epidemiologie,  
Liebigstraße 27, 04103 Leipzig, Telefon: (0341) 971 61 00, Telefax:  
(0341) 971 61 09, e-mail: Loeffler@imise.uni-leipzig.de. Verantwortlich für die  
Mitteilungen der Deutschen Gesellschaft für Medizinische Informatik, Biometrie  
und Epidemiologie e. V.: Dipl.-Volksw. Th. Banasiewicz, Herbert-Lewin-Straße 1,  
50931 Köln, Tel. (0221) 4004-865.

**Verlag:** Vertrieb und Werbung: Urban & Fischer Verlag GmbH & Co. KG,  
Niederlassung Jena, PF 100537, D-07705 Jena; Tel. (03641) 626-3; Fax  
(03641) 62 65 00; e-mail: journals@urbanfischer.de. Herstellung und Anzei-  
gen: Verlag Eugen Ulmer GmbH & Co., Wollgrasweg 41, 70599 Stuttgart, Tel.  
(07 11) 45 07-0; e-mail: info@ulmer.de. Postscheckkonto Stuttgart 74 63-700,  
Zürich 80-47072, Wien 1083.662 Deutsche Bank AG, Stuttgart, Kto. 14/76 878,  
Südwestbank AG, Stuttgart, Kto. 741 371 006, Herstellung: Sigrid Wolf, Tel.  
(07 11) 45 07-194. Verantwortlich für die Anzeigen: Dieter Boger, Lilien Sever,  
Tel. (07 11) 45 07-144, z. Z. ist die Anzeigenpreisliste Nr. 10 gültig. Anzeigen-  
schluß: am 20. der Monate Januar, April, Juli, Oktober.

**Druck:** Druckhaus »Thomas Müntzer« GmbH, Neustädter Straße 1-4,  
99947 Bad Langensalza, Telefon (036 03) 3 99-0.

**Abonnementsverwaltung:** SFG-Servicecenter Fachverlage GmbH, Zeitschriften-  
vertrieb: Barbara Dressler, Villengang 2 D-07745 Jena; Telefon: (036 41)  
62 64 44; Fax (036 41) 62 64 43.

**Bezugshinweise:** Das Abonnement gilt bis auf Widerruf oder wird auf Wunsch  
befristet. Die Lieferung der Zeitschrift läuft weiter, wenn sie nicht bis zum  
31. 10. eines Jahres abbestellt wird.

**Erscheinungsweise:** (1999) 1 Jahrgang mit 4 Hefen.

**Abo-Preise:** (1999) 1 Jahrgang 392,- DM; Einzelheftpreis 110,- DM. Alle  
Preisangaben verstehen sich zuzüglich Versandkosten.

Vorzugspreis für Mitglieder der Deutschen Region der Internationalen Bio-  
metrischen Gesellschaft auf Anfrage beim Verlag.

Folgende Kreditkarten werden zur Zahlung akzeptiert: Visa/Eurocard/Master-  
card/American Express (bitte Kartennummer und Gültigkeitsdauer angeben).

**Bankverbindung:** Deutsche Bank AG Jena, Konto-Nr. 6 284 707, BLZ  
820 700 00.

**Indexed in:** »Biological Abstracts« und »Current Index to Statistics«.

**Copyright:** Die in der Zeitschrift veröffentlichten Beiträge sind urheberrechtlich  
geschützt: 1.) Mit der Abgabe des Manuskripts versichert der Autor, daß er  
allein befugt ist, über die urheberrechtlichen Nutzungsrechte an seinem Bei-

trag, einschließlich eventueller Bild- und anderer Reproduktionsvorlagen zu  
verfügen und daß der Beitrag keine Rechte Dritter verletzt. 2.) In Erweiterung  
von § 38 Abs. 1 UrhG räumt der Autor dem Verlag für die Dauer des Urheber-  
rechts das räumlich und mengenmäßig unbeschränkte Recht der Vervielfäl-  
tigung und Verbreitung (Verlagsrecht) beziehungsweise der unkörperlichen  
Wiedergabe der Beiträge ein, auch zur Verwertung außerhalb der Zeitschrift,  
für die er ursprünglich bestimmt war. Die Übertragung erfolgt auf die Dauer von  
fünf Jahren ausschließlich. 3.) Im Rahmen von Ziffer 2 räumt der Autor dem  
Verlag ferner die ausschließlichen Nutzungsrechte am Beitrag ein, so a) das  
Recht der Übersetzung in Fremdsprachen, das Recht zum ganzen oder teil-  
weisen Vorabdruck und Nachdruck – auch in Form eines Sonderdrucks –, zur  
Übersetzung in andere Sprachen, zu sonstiger Bearbeitung und zur Erstellung  
von Zusammenfassungen (Abstracts), b) das Recht zur Veröffentlichung einer  
Mikrokopie, Microfiche- und Microformausgabe, zur Nutzung im Wege von  
Bildschirmtext, Videotext und ähnlichen Verfahren, zur Aufzeichnung auf Bild-  
und/oder Tonträger und zu deren öffentlicher Wiedergabe, c) das Recht zur  
maschinenlesbaren Erfassung und elektronischen Speicherung auf einem  
Datenträger und in einer eigenen oder fremden Online-Datenbank, das Recht  
Teile des Beitrages (Beitragskopf mit Zusammenfassungen und Schlüsselwör-  
te in deutsch und englisch) ins Internet zu stellen, zum Download in einem  
eigenen oder fremden Rechner zur Wiedergabe am Bildschirm, sei es unmit-  
telbar oder im Wege der Datenfernübertragung, sowie zur Bereithaltung in  
einer eigenen oder fremden Online-Datenbank zur Nutzung durch Dritte, d) das  
Recht zu sonstiger Vervielfältigung, insbesondere durch fotomechanische oder  
ähnliche Verfahren und zur Nutzung im Rahmen eines sogenannten Kopien-  
versandes auf Bestellung, e) das Recht zur Vergabe der vorgenannten Nut-  
zungsrechte an Dritte im In- und Ausland sowie die von der Verwertungsgesell-  
schaft WORT wahrgenommenen Rechte einschließlich der entsprechenden  
Vergütungsansprüche. Der Verlag wird über die Rechte gemäß Punkt 2 und  
3a nur mit Zustimmung des Autors verfügen und sich um eine angemessene  
Honorierung bemühen. Fotokopien für den persönlichen Gebrauch dürfen nur  
von den einzelnen Beiträgen oder Teilen daraus als einzelne Kopien erstellt  
werden.

Printed in Germany  
© 1999 Verlag Eugen Ulmer Stuttgart; Urban & Fischer Verlag Jena



# Modellbasierte Prädiktion der Diabetesprävalenz

## Model-based prediction of diabetes prevalence

P. Heinke, E. Salzsieder

### Zusammenfassung

*Eine zielgerichtete und effiziente Planung von Maßnahmen zur Verbesserung der Gesundheitsfürsorge des Diabetes mellitus erfordert sowohl aktuelle Daten zur Häufigkeit in der Bevölkerung als auch differenziertere Angaben zu den diabetesspezifischen Komplikationsraten in Abhängigkeit vom Alter und von der Krankheitsdauer. Dazu wird ein Modell zur Vorhersage der Diabetikerzahlen entwickelt, welches unter Berücksichtigung der bei Diabetikern verkürzten Lebenserwartung Aussagen zur Anzahl der Betroffenen und zu deren Altersstruktur erlaubt. Durch Variation der einzelnen Eingangsgrößen können mittels dieses Modells auch Aussagen zu möglichen Ursachen von Bestandsveränderungen gemacht werden.*

*Mit dem Modell wird ein Anstieg der Anzahl der Diabetiker für die östlichen Bundesländer der BRD von 689000 im Jahr 1989 auf 826600 im Jahr 2000, d.h. ein Zuwachs von 20,0%, prognostiziert. Als Ursachen dafür werden Verschiebungen in der Bevölkerungsstruktur, Veränderungen in der Inzidenz und lebensverlängernde Maßnahmen identifiziert.*

### Schlüsselwörter

*Epidemiologie, Diabetes mellitus, Lebenserwartung, Vorhersagemodell, Überlebensmodell*

### Summary

*Planning measures to improve health care requires age-related epidemiological data of diabetes mellitus and its complications. Therefore we developed a model-based procedure to predict the number of diabetic patients within the entire population. The model permits to estimate prevalence rates with respect to the reduced life expectancy of the diabetic population. By separate variation of the input variables of the model it is also possible to identify causes of alterations in the prevalence of diabetes mellitus.*

*It is predicted, that we have to expect an increase in the number of diabetic patients for the Easter part of Germany (former GDR) by 20.0%, i.e. from 689,000 in the year 1989 up to 826,600 in the year 2000. This increase is mainly caused by age-related structural changes in the general population, by increasing incidence rates of the disease and by measures aimed to increase life expectancy of diabetic patients.*

**Key words**

*epidemiology, diabetes mellitus, life expectancy, model-based prediction, survival model*

**1 Einführung**

Die demographische Entwicklung in den Industrieländern Westeuropas ist durch eine zunehmende Alterung der Bevölkerung gekennzeichnet. Damit rücken vor allem die in den höheren Altersklassen auftretenden chronischen Erkrankungen, darunter der Diabetes mellitus, aus Sicht der Gesundheitsfürsorge weiter in den Vordergrund. Für Aussagen über den materiellen Aufwand bei der Behandlung der Krankheit, zur Planung medizinischer Forschung und von Programmen zur Verbesserung der Gesundheitsfürsorge sind epidemiologische Maßzahlen unerlässlich. Von besonderem Interesse sind hierbei Voraussagen über die zukünftig zu erwartende Entwicklung der Patientenzahlen.

Dazu sind bereits verschiedene Methoden und Modellansätze erarbeitet worden. Die Standardisierungsmethode überträgt die gegenwärtigen altersspezifischen Prävalenzraten auf zukünftige Bevölkerungsstrukturen (HELMS, 1992). Regressionsmodelle beschreiben die zeitliche Abhängigkeit bekannter epidemiologischer Parameter durch Funktionen, die dann in den zu untersuchenden Zeitbereich extrapoliert werden können. Stochastische Modelle basieren in der Regel auf der strukturierten Beschreibung von unterschiedlichen Gruppen von Betroffenen und nicht Betroffenen (z. B. Nichtdiabetiker, Typ1- und Typ2-Diabetiker). Ausgegangen wird hierbei zumeist von einer bekannten Ausgangsverteilung und den jährlichen Übergangswahrscheinlichkeiten zwischen den Gruppen. Sind diese bekannt, läßt sich das Gesamtsystem Jahr für Jahr weiterverfolgen (HERMAN, 1984, HAUSER, 1988, GREEN, 1996). Ein anderer Ansatz ist das Überlebensmodell (RUWAARD, 1993), bei dem das System durch die Verweildauer, d.h. durch die Überlebenskurve seiner Elemente beschrieben wird.

Beim Überlebensmodell wird auf der Grundlage von gesicherten epidemiologischen Studien davon ausgegangen, daß die Diabetiker bei Manifestation, d.h. bei Diagnosestellung ihrer Krankheit, eine im Vergleich zur übrigen Bevölkerung verminderte Überlebenswahrscheinlichkeit und damit eine reduzierte alters- und geschlechtsabhängige fernere Lebenserwartung (fle) haben. Zur Beurteilung der Verminderung der Lebenserwartung liegen hinsichtlich Untersuchungsgruppen, Design, Altersgruppen, Auswertung, Vergleichsgruppen, Zeitraum und Zeitdauer sehr verschiedenartige Untersuchungen vor, die aufgrund unterschiedlicher methodischer Ansätze nur schlecht vergleichbar sind. Als Reduzierung der Lebenserwartung wird jedoch übereinstimmend die Differenz zwischen den ferneren Lebenserwartungen von Bevölkerung und Diabetikern definiert.

Auf der Grundlage der Verweildauer einzelner Bevölkerungsgruppen im demographischen Gesamtsystem eines definierten Einzugsgebietes wird nachfolgend ein Modellansatz vorgeschlagen, mit dem die Diabetikerzahlen in den östlichen Bundesländern der BRD (Gebiet der ehemaligen DDR) für das Jahr 2000 prognostiziert und mögliche Ursachen für deren Zunahme identifiziert werden sollen.

Zur Überprüfung des Modells dienen die Daten des Zentralen Diabetesregisters der früheren DDR (MICHAELIS, 1991). Von 1960 bis 1989 wurden durch alle 227 Kreis- bzw. Stadtbezirksstellen der Diabetikerdispensaries zum Jahresende der Bestand und die im Laufe des Jahres neuerkrankten und gestorbenen Diabetiker erfaßt, die dann durch das Zentralinstitut für Diabetes in Karlsburg aggregiert wurden. Die Angaben erfolgten aufgeschlüsselt nach Alter (bis 1981 in Dekaden, danach in Pentaden), Geschlecht und der Art der Behandlung. Daneben wurden summarisch die Anzahl der Remissionen sowie die der Zu- und Abgänge durch Ortsveränderung registriert.

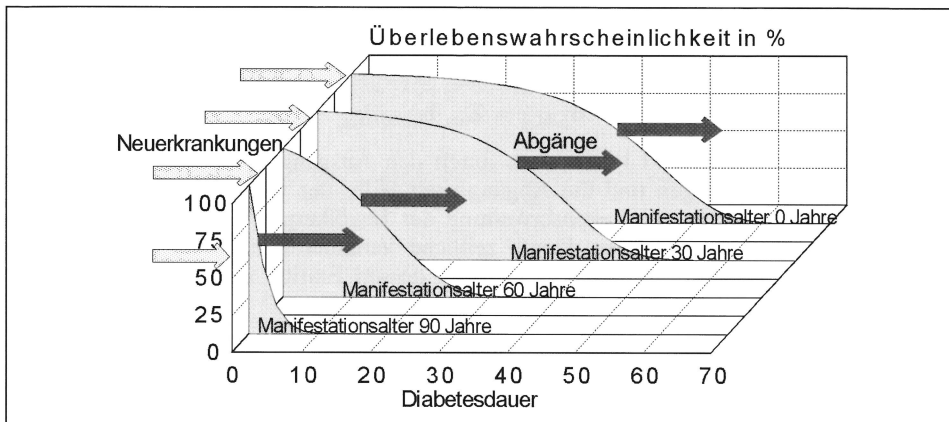


Abbildung 1: Überlebensmodell: System der Diabetiker mit den jährlich neu diagnostizierten Krankheitsfällen als Input und den entsprechend ihrer alters-, geschlechts- und krankheitsdauerabhängigen Sterbewahrscheinlichkeit das System verlassenden diabetischen Patienten als Output

## 2 Modellbeschreibung

### 2.1 Allgemeine Beschreibung

Betrachtet wird das System der Diabetiker (Abbildung 1), in das als Eingangsgröße die jährlich diagnostizierten Neuerkrankungen eingehen. Diese haben eine vom Geschlecht, dem Alter bei der Krankheitsfeststellung (Manifestationsalter) und der Krankheitsdauer (Diabetesdauer) abhängige Überlebenswahrscheinlichkeit, die bei Diabetikern im Vergleich zur übrigen Bevölkerung vermindert ist. Entsprechend ihrer Überlebenskurve verläßt jedes Jahr eine bestimmte Anzahl von Diabetikern das betrachtete System.

Der Abgang entsteht durch die gestorbenen Diabetiker. Unter geeigneten Annahmen über den Anfangsbestand, die Neuerkrankungen und die Verweildauer im System ist es dann möglich, das System zeitbezogen zu beschreiben und damit epidemiologische Maßzahlen zur bevölkerungsbezogenen Diabetesentwicklung zu prognostizieren.

### 2.2 Mathematische Beschreibung

Zu Beginn der Prognose, im Jahr  $j = j_0$ , liegt der Bestand an Diabetikern  $b_{j_0, ma, dd, g}$  nach dem Manifestationsalter  $ma$ , der Diabetesdauer  $dd$  und dem Geschlecht  $g$  gegliedert vor. Zu diesen kommen im nachfolgenden Jahr  $j + 1$  die neu diagnostizierten Krankheitsfälle

$$n_{j+1, ma, dd, g} \quad \text{mit} \quad dd = 0$$

hinzu. Gleichzeitig verringert sich entsprechend der Überlebenswahrscheinlichkeit für  $j > j_0$  der vorhandene Bestand. Der Einbezug des Sterbegeschehens erfolgt durch Multiplikation der alten Bestände mit den zugehörigen aktuellen Überlebenswahrscheinlichkeiten  $\bar{u}_w$ . Damit ergibt sich der neue Bestand im Jahr  $j + 1$  zu

$$B_{j+1} = \sum n_{j+1, ma, dd, g} + \sum b_{j, ma, dd, g} \cdot \bar{u}_{w_{j+1, ma, dd, g}}.$$

Da der nach dem Manifestationsalter, der Diabetesdauer und damit auch nach dem Alter gegliederte Bestand in Form einer Matrix dargestellt wird, ist eine Aggregation nach Alter, Diabetesdauer und Manifestationsalter leicht möglich.

$$\text{Alter:} \quad B_{j,a,g} = \sum_a b_{j,a,ma,dd,g} \quad \text{mit} \quad a = ma + dd$$

$$\text{Manifestationsalter:} \quad B_{j,ma,g} = \sum_{dd} b_{j,ma,dd,g}$$

$$\text{Diabetesdauer:} \quad B_{j,dd,g} = \sum_{ma} b_{j,ma,dd,g}$$

Das betrachtete System wird beschrieben durch den Ausgangsbestand, die Zugänge in Form der Neuerkrankungen und die Abgänge mit Hilfe der Überlebenswahrscheinlichkeit, die durch die Parameter Lebenserwartung der Bevölkerung und dessen Reduktion bei den Diabetikern festgelegt ist. Durch zeitliche Veränderungen eines Parameters bei gleichzeitigem Konstanthalten der übrigen kann dessen Einfluß auf die Prognoseergebnisse analysiert werden.

### 3 Datengewinnung

#### 3.1 Bestimmung der Reduktion der Lebenserwartung bei Diabetikern (rle)

Aus den Longitudinalstudien von Deckert (1978), Goodkin (1975), Marks (1971), Panzram (1981), Schneider (1994) und Thölke (1994) sind die problembezogenen Daten entnommen und die Parameter der reduzierten Lebenserwartung vermöge einer altersabhängigen Exponentialfunktion ( $rle = a \cdot \exp(-c \cdot \text{alter})$ ) mittels nichtlinearer Regression bestimmt worden (Abbildung 2). Die großen Schwankungen in den jüngeren Altersklassen basieren auf vergleichsweise kleinen Fallzahlen in diesem Altersbereich.

Da in einigen Studien statt der fernerer die medianen Lebenserwartungen (mle) verwendet werden, die wegen der Schiefe der Verteilung der Sterbefälle in den jüngeren Altersklassen größer, in den älteren kleiner als die fle sind, wurde die rle anhand der vollständigen allgemeinen Sterbetafel für die DDR (Statistisches Jahrbuch 1990) korrigiert. Die gleichzeitig in Abhängigkeit von den unterschiedlichen Lebenserwartungen der beiden Geschlechter verwendeten Werte sind der Tabelle 1 zu entnehmen.

Seit der Einführung des Insulins im Jahre 1922 stieg die Lebenserwartung der Diabetiker stetig (JOSLIN, 1952) und im Vergleich zur Bevölkerung in höherem Maße an. Dieses

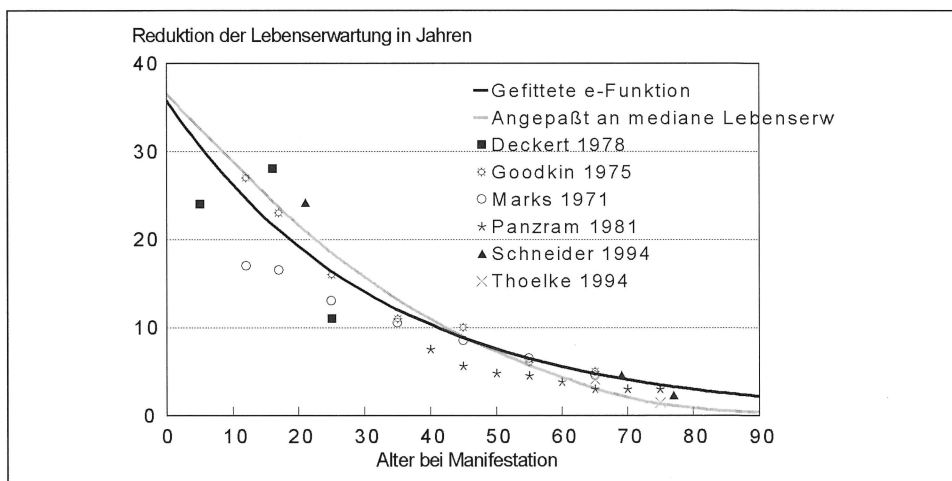


Abbildung 2: Reduktion der Lebenserwartung für diabetische Patienten im Vergleich zur Bevölkerung in Abhängigkeit vom Alter bei der Krankheitsfeststellung

Tabelle 1: Fernere Lebenserwartung der Bevölkerung (fle), deren Reduktion bei Diabetikern gegenüber der Bevölkerung (rle) und das „angepaßte“ Alter der Diabetiker bei Manifestation in Jahren

Alter	Lebenserwartung der Bevölkerung (fle)		Reduktion der Lebenserwartung bei Diabetikern (rle)			„angepaßtes“ Alter der Diabetiker bei Manifestation*	
	männlich	weiblich	e-Funktion	korrigiert		männlich	weiblich
0	72,3	78,2	35,8	36,0	36,0	37,4	38,1
10	62,3	68,3	26,3	29,0	31,0	40,5	42,2
20	52,4	58,3	19,3	20,0	23,4	41,5	44,6
30	42,5	48,3	14,1	14,6	17,0	46,3	48,3
40	32,9	38,4	10,4	10,4	11,6	52,3	53,0
50	23,6	28,6	7,6	7,0	7,8	59,4	59,4
60	15,4	19,4	5,6	4,2	4,6	67,0	66,4
70	8,8	11,3	4,1	1,8	2,4	74,5	74,6
80	4,4	5,5	3,0	0,8	1,0	83,6	83,4
90	2,0	2,2	2,2	0,4	0,4	93,7	93,0

\* entspricht dem Alter der Bevölkerung hinsichtlich der Überlebenskurve bei der Manifestation

ist gleichbedeutend mit einer entsprechenden Verminderung der rle. Da die verwendeten Studien im Mittel in den 70er Jahren begannen, wurde das Jahr 1970 als Bezugspunkt für die so bestimmte rle genommen.

### 3.2 Bestimmung der Überlebenskurve der Diabetiker

Aus der vollständigen allgemeinen Sterbetafel der DDR für 1988/89 (Statistisches Jahrbuch 1990) sind die aus einer Kohorte von je 100 000 Personen in den einzelnen Altersgruppen Überlebenden entnommen worden. An diese Daten wurde mittels nichtlinearer Regression eine Gompertz-Funktion der Form

$$y = f(100 - a) = b_1 \cdot \exp(b_2 \cdot \exp(b_3 \cdot (100 - a)))$$

angepaßt, wobei y die Anzahl der Überlebenden der Kohorte,  $a < 100$  das Alter in Jahren und  $b_1$ ,  $b_2$ ,  $b_3$  die Funktionsparameter darstellen. Die Anpassung ergab folgende Punktschätzer:

$$\text{männlich } b_1 = 98642 \quad b_2 = -8,03972 \quad b_3 = -0,0918117 \quad r = 0,99995$$

$$\text{weiblich } b_1 = 98617 \quad b_2 = -6,79608 \quad b_3 = -0,1099800 \quad r = 0,99999$$

Mit diesen Parametern ist der Absterbeverlauf der 0-jährigen festgelegt. Bezieht man die nachfolgenden Werte auf den Funktionswert des Alters  $a_1$ , so erhält man die Überlebenswahrscheinlichkeit eines  $a_1$ -jährigen (Abbildung 3)

$$y_{a_1} = f(100 - a)/f(100 - a_1)$$

mit  $a > a_1$ .

Aus der Funktion ist die mediane Lebenserwartung mle, d.h. die Zeit, nach der noch die Hälfte der Kohorte lebt, zu bestimmen (Auflösen der Gleichung  $\frac{1}{2} = f(100 - a_1 - mle)/f(100 - a_1)$  nach mle). Diese ist für die Diabetiker um rle vermindert. Damit liegen 2 Punkte der Überlebensfunktion fest, der eine beim Alter der Krankheitsfeststellung  $a = ma$ , bei dem noch alle Personen leben (100%), und der andere bei  $a = ma + mle$ , bei dem noch 50% der Personen leben. Zur Bestimmung des übrigen Kurvenverlaufes

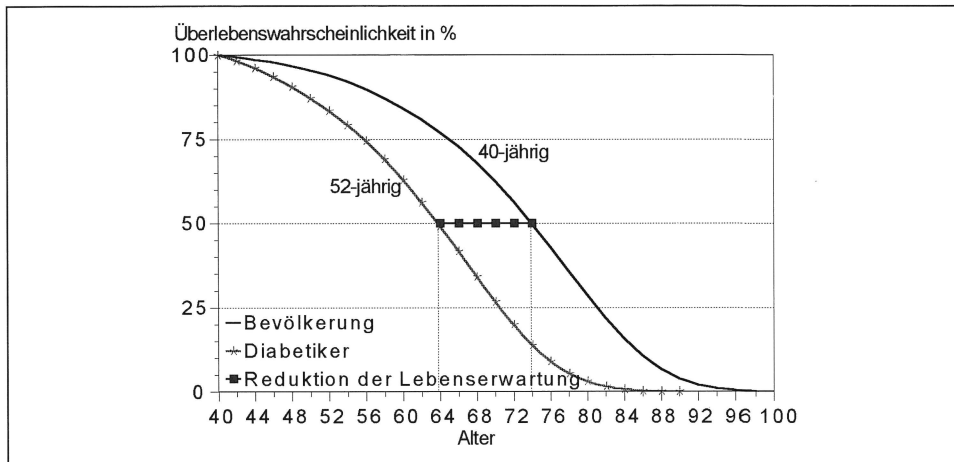


Abbildung 3: Überlebenskurve eines 40-jährigen der Bevölkerung im Vergleich zu einem im Alter von 40 Jahren an Diabetes Erkrankten, dessen Lebenserwartung um 10,4 Jahre vermindert ist und dessen Überlebenskurve damit dem eines 52-jährigen der Bevölkerung entspricht

wird die Annahme gemacht, daß die Absterbefunktionen von Bevölkerung und Diabetikern gleich sind, daß sie aber bei den Diabetikern nicht mehr dem Alter  $ma$  entsprechen, sondern zu einem höheren Alter  $ma + a'$  verschoben sind, bei der die  $mle$  um  $rle$  reduziert ist. Das bedeutet die Lösung von

$$\frac{1}{2} = f(100 - (ma + a') - (mle - rle)) / f(100 - (ma + a'))$$

nach  $ma + a'$ , was analytisch keine Schwierigkeiten bedeutet. Das so gefundene „angepaßte“ Manifestationsalter bei Diabetikern ist der Tabelle 1 zu entnehmen.

### 3.3 Rahmenbedingungen für die Modellentwicklung

Zur Evaluierung des Modells und zur eigentlichen Vorhersage der bevölkerungsbezogenen Entwicklung der Diabeteshäufigkeit werden neben den Neuerkrankungen die Bestände an Diabetikern aus dem Zentralen Diabetesregister der DDR verwendet (MICHAELIS, 1991). Diese Angaben liegen im Gegensatz zu den Lebenserwartungen nur in 10- bzw. 5-Jahresklassen vor, so daß vorher der Übergang auf 1-Jahresklassen notwendig war. Dazu wurden die über die Altersklassen gegebenen Treppenfunktionen der Prävalenz und Inzidenz so interpoliert, daß in den Altersklassen die gemessenen Anzahlen von Personen und das Integral der zugehörigen Lebenserwartung übereinstimmen.

Für die Vorhersage sind die bis 1989 vorhandenen Inzidenzen mittels logistischer Regression bis zum Jahr 2000 extrapoliert worden. Der Übergang auf die entsprechenden absoluten Neuerkrankungen erfolgte über die anhand der Sterbewahrscheinlichkeiten der Allgemeinen Sterbetafel für die DDR von 1988/89 (Statistisches Jahrbuch, 1990) fortgeschriebene Bevölkerungsstruktur von 1995 [Mitteilung des Statistischen Bundesamtes], wobei eine jährliche Zunahme der Geburtenzahlen von 2% angenommen wurde.

Die zeitliche Entwicklung der Lebenserwartung der Bevölkerung und der Reduktion der Lebenserwartung bei den Diabetikern  $rle$  wird über entsprechende Korrekturen  $dle$  und  $drle$  berücksichtigt. Damit ergibt sich für einen im Alter  $ma$  an Diabetes Erkrankten der Bezugspunkt  $bz$  in der Überlebenskurve für das Jahr  $j$  zu

$$bz_{j,ma} = ama_{ma} + dle_j + drle_j$$



und die Überlebenswahrscheinlichkeit zwischen den Zeitpunkten  $a_1$  und  $a_2$  zu

$$\text{ülw} = y(100 - bz_{j, \text{ma}, g} - a_2) / y(100 - bz_{j, \text{ma}, g} - a_1)$$

mit  $a_2 > a_1$ ,  $a_{\text{ma}}$  = "angepaßtes" Manifestationsalter,  $bz$  = zeitlich korrigierter Bezugspunkt (Alter).

Die Verweildauer der Diabetiker im System ist abhängig von Manifestationsalter und Diabetesdauer. Da keine entsprechenden Daten dafür vorliegen, ist es notwendig, den Anfangsbestand entsprechend aufzuteilen. Näheres dazu ist dem Anhang zu entnehmen.

## 4. Modellanwendung

### 4.1 Historische Validierung des Modells

Zur Überprüfung des Modells wurde eine retrospektive Analyse über 25 Jahre für den Zeitraum von 1964 bis 1989 durchgeführt.

Die modellgestützte historische Simulation zeigte am Ende des Vorhersagezeitraumes eine um 3100 (weiblich +2860, männlich +240) zu hohe Anzahl der modellgestützt geschätzten Diabetikerzahlen im Vergleich zu den tatsächlich im Diabetesregister erfaßten Fallzahlen. Diese Differenz entspricht einem Vorhersagefehler des Modells von 0,4% und ist praktisch bedeutungslos.

In den bestandsmäßig größten Altersklassen 50 J bis 90 J lagen die Abweichungen zwischen -7,2% und 6,0%, während in den übrigen Klassen auf Grund der geringeren Fallzahlen und der dadurch bedingten größeren Streuungen die Abweichungen erwartungsgemäß entsprechend größer waren (-0,8% bis 10,1%).

### 4.2 Vorhersage der Diabetikerzahlen

Mittels prospektiver Vorhersagen werden ausgehend vom Jahr 1989 entsprechende Werte für das Jahr 2000 geschätzt. Dabei wird angenommen, daß die das Modell beeinflussen Parameter Inzidenz, Lebenserwartung der Bevölkerung und Reduktion der Lebenserwartung der Diabetiker dem gleichen Trend wie in den Jahren vor 1989 unterworfen sind.

Mit dem Modell wird für den Zeitraum von 1989 bis 2000 eine weitere Zunahme der Diabetikerpopulation um 20,0% prognostiziert, was dann einen Bestand von 826 600 Diabetikern für die östlichen Bundesländer der BRD erwarten läßt (Tabelle 2). Der Zuwachs wird dabei vorwiegend durch die männlichen Diabetiker verursacht (männlich +40,9% vs. weiblich +7,3%).

Die Geschlechtsunterschiede resultieren vorwiegend aus dem Wirken demographischer und diabetogener Einflüsse auf die Prognosewerte (männlich +34,3%, weiblich +5,7%). Dazu zählen das demographische Ungleichgewicht, das sich in einem Übergewicht der

Tabelle 2: Vorhersage der Anzahl der Diabetiker für das Jahr 2000

	1989 Anzahl	%	2000 Anzahl	%
männlich	260 000	3,3	373 000	5,2
weiblich	429 100	5,0	450 700	5,9
Gesamt	689 100	4,2	824 000	5,6

Tabelle 3: Anteil der einzelnen Einflußfaktoren auf die Zunahme des Diabetikerbestandes im Jahr 2000

Einflußfaktor	absolut		Gesamt	relativ		Gesamt
	männlich	weiblich		männlich	weiblich	
Zunahme 1989–2000	106 400	31 200	137 600	40,92	7,27	19,97
Änderungen der Bevölkerungsstruktur	89 130	24 560	113 690	34,28	5,72	16,50
Änderungen der Inzidenz	16 630	4 410	21 040	6,40	1,03	3,05
Lebenserwartung der Bevölkerung	540	2 140	2 680	0,21	0,50	0,39
Allgemeine Verminderung der rle	100	90	190	0,04	0,02	0,03

Zu- über die Abgänge im System der Diabetiker bei konstanter Verweildauer zeigt, die natürliche Bewegung der Bevölkerung, die zu einem Umbau der Geschlechts- und Altersstruktur und damit auch zu höheren Neuerkrankenzahlen führt, und die geschlechts- und altersabhängigen Inzidenzveränderungen (+6,4% bzw. 1,0%). Demgegenüber haben die die Verweilzeit im System beeinflussenden Parameter nur einen geringen Einfluß (höhere Lebenserwartung der Bevölkerung +0,2 bzw. +0,5%, allgemeine Reduktion der Lebenserwartung bei Diabetikern +0,04 bzw. +0,02%) (Tabelle 3).

## 5 Diskussion

Zur Vorhersage zukünftiger Diabetikerzahlen wurde das bekannte Modell von RUWAARD [1993] problemorientiert modifiziert. Dabei ist die Verweilzeit der Diabetiker im betrachteten System nicht durch einen Schnitt bei der altersentsprechenden Lebenserwartung, sondern durch Einbezug der Überlebenskurven integriert worden. Wegen der bekannten Asymmetrie der Sterbefälle, deren Verteilungen mit wachsendem Manifestationsalter von links- nach rechtsschief wechseln und die sich in unterschiedlichen Werten für die fernere und mediane Lebenserwartung zeigen, entspricht dieses Vorgehen eher den tatsächlichen Sterblichkeitsverhältnissen. In die gleiche Richtung läuft die Verwendung alters- und geschlechtsspezifischer Werte für die Reduktion der Lebenserwartung bei Diabetikern. Obwohl dafür aus der Literatur keine Angaben vorliegen, scheint die Annahme plausibel, daß bei den Frauen im Vergleich zu den Männern mit der z.T. erheblich höheren Lebenserwartung auch eine erhöhte Reduktion der Lebenserwartung einhergeht. Für die Gültigkeit dieser Annahmen sprechen die sehr guten Vorhersageergebnisse der hier durchgeführten historischen Analyse. Die Abweichung zwischen vorhandenem Diabetikerbestand und Vorhersage in der historischen Validierung betrug bei uns 0,4% im Vergleich zu 10,0% bei Ruwaard. Die deutliche Verbesserung der Ergebnisse ist vor allem auf die Verwendung der Daten des Diabetesregisters der ehemaligen DDR, die Einführung der altersabhängigen Überlebenskurven und der zeitabhängigen Reduktion der Lebenserwartung bei den Diabetikern zurückzuführen. Mit dem Register liegt ein einmaliges Datenmaterial zur Epidemiologie des Diabetes auf Bevölkerungsebene über 30 Jahre vor, das es uns erlaubte, das Vorhersagemodell in dieser Form aufzustellen und retrospektiv zu validieren. Dadurch wurde es überhaupt erst möglich, differenziert die Auswirkungen von Einflußfaktoren auf die Bestandserhöhung zu untersuchen.

Im vorgestellten Modell sind verschiedene Parameter implementiert, die die Vorhersageergebnisse beeinflussen. Dazu zählen demographische Daten (Größe und Struktur des Ausgangsbestandes und die daraus resultierende Anzahl von Neuerkrankungen), die die Verweildauer beeinflussenden Größen (Lebenserwartung der Bevölkerung, Reduktion der

Lebenserwartung der Diabetiker im Vergleich zur Bevölkerung, Überlebenskurve) sowie Änderungen des individuellen Erkrankungsrisikos (Inzidenz). Mit wachsendem Vorhersagezeitraum nimmt dabei der Einfluß des Ausgangsbestandes wegen des großen Anteils der relativ kurzlebigen Altersdiabetiker schnell immer weiter ab. Größere Auswirkungen haben dagegen Änderungen bei den Neuerkrankungen und der Verweildauer im System, die nachhaltig die Bestandsgröße verändern.

Es ist davon auszugehen, daß zeitliche Veränderungen in den Eingangsparametern weder sprunghaft erfolgen noch in allen Altersklassen gleich groß sind. Daher dürften Simulationsergebnisse, bei denen der Einfluß solcher sprunghaft konstanter Änderungen zugrunde gelegt wird (HAUSER, 1988, RUWAARD, 1993, JUTZI, 1994), kaum der Realität entsprechen.

Eine wichtige Frage ist die nach bestandsverändernden Faktoren. Mit dem Modell ist es nicht möglich, differenzierte Aussagen zu den in der Literatur behandelten Ursachen (genetische Faktoren, Umwelt, Lebensstil, Ernährung, Stress usw.) (ZIMNET, 1982) zu machen, da in das Modell nur Parameter aufgenommen worden sind, die das Ergebnis deren Wirkens sind, d.h. solche Parameter, die sich letztendlich in Inzidenz- und Prävalenzraten sowie in der Lebensdauer widerspiegeln. Durch deren Variation wird es mittels des hier beschriebenen Modells aber erstmals möglich, deren separate Auswirkungen auf die Entwicklung des Diabetikerbestandes abschätzen.

In Übereinstimmung mit epidemiologischen Studien anderer Autoren (MICHAELIS, 1991, JUTZI, 1994) weist der mittels des beschriebenen Modellansatzes ermittelte Bestand an Diabetikern in den östlichen Bundesländern der BRD eine steigende Tendenz auf. Dabei verhalten sich beide Geschlechter unterschiedlich. Der weibliche Diabetikerbestand wächst langsam (+2800 Fälle jährlich), d.h. im System befinden sich Zu- und Abgänge bzw. inzidente und gestorbene Fälle, fast in einem Gleichgewicht. Bei den Männern dagegen nimmt der Bestand jährlich im Durchschnitt um 9670 Fälle zu. Die Erklärung dafür liegt in einer großen Umstrukturierung des Bevölkerungsbaumes. Dominierender Einflußfaktor ist die demographische Entwicklung der Bevölkerung, der sich in einem Umbau der Alters- und Geschlechtsstruktur manifestiert. Dabei werden vor allem die durch die Weltkriege reduzierten, jetzt älteren und damit den Diabetikerbestand bestimmenden Altersklassen bei den Männern durch die nachfolgenden und gegenüber dem weiblichen Geschlecht überwiegenden starken Jahrgänge aufgefüllt. Diese natürliche Bevölkerungsbewegung erzeugt damit ein starkes Anwachsen der Neuerkrankungen an Diabetes, welches sich in einem weiteren Verstärken des demographischen Ungleichgewichts niederschlägt. Dagegen hat die anhaltende Halbierung der Neugeborenenzahlen gegenüber 1989 noch keinen Einfluß auf den absoluten Bestand an Diabetikern im Jahre 2000. Die starke Reduktion der eigentlich bestandsmäßig größten Jahrgänge hingegen macht sich mit einem entsprechenden Prävalenzanstieg von 4,2 auf 5,5% bemerkbar.

Zusammenfassend läßt sich feststellen, daß das Überlebensmodell in der vorgestellten Form gut geeignet ist, sowohl Vorhersagen zu der weiteren Entwicklung des bevölkerungsbezogenen Diabetikerbestandes zu beschreiben als auch Aussagen über die Ursachen von Bestandsveränderungen zu gewinnen. Die Resultate dürften somit auch von nachhaltigem Interesse für ein breit angelegtes Qualitätsmanagement in der Diabetikerbetreuung sowie für die gezielte Ressourcenplanung in der Gesundheitsfürsorge für die betroffene Bevölkerungsgruppe sein. Die Modellrechnung erfordert nur wenige Eingangsparameter (Anfangsbestand, Überlebenskurve der Diabetiker, jährliche Neuerkrankungen). Durch die Berücksichtigung von alters- und geschlechtsspezifischen Trends bei den Inzidenzen und der Lebenserwartung ist das wirkliche Geschehen mittels des Modells sehr sensitiv abgebildet worden. Für die Erweiterung des Vorhersagezeitraumes erscheint es erforderlich, die Eingangsparameter und die Ergebnisse der Modellrechnung in geeigneten Zeitabständen erneut zu validieren.

## 6 Literatur

- DECKERT, T., POULSEN, J. E., LARSEN, M. (1978): Prognosis of diabetics with diabetes onset before the age of thirtyone. *Diabetologia* **14**, 363–370.
- GOODKIN, G. (1975): Mortality factors in diabetes. *J. Occupat. Med.* **17**, 716–721.
- GREEN, A., SJOLIE, A. K., ESHOJ, O. (1996): Trends in the epidemiology of IDDM during 1970–2020 in Fyn county, Denmark. *Diabetes Care* **19**, 801–806.
- HAUSER, F., ANDEL, M. (1988): Projection of the prevalence of type 1 and type 2 diabetes. In: *Advances in biomedical measurement*, Carson, E. R., Kneppo, P., Krekule, I. (Eds.). New York: Plenum Press, 407–413.
- HELMS, R. B. (1992): Implications of population growth on prevalence of Diabetes. A look at the future. *Diabetes Care* **15**, Suppl. 1, 6–9.
- HERMAN, W. H., SINNOCK, P., BRENNER, E. (1984): An epidemiological model for Diabetes mellitus: Incidence, prevalence, and mortality. *Diabetes Care* **7**, 367–371.
- JOSLIN, E. P. (1952): The duration of life and the prognosis in diabetes. In: *The treatment of Diabetes mellitus*, Joslin, E. P., Root, H. F., White, P., Marble, A. (Eds.). Philadelphia: Lea & Feabiger, 261–272.
- JUTZI, E. (1994): Wie häufig ist Diabetes in Deutschland und wie häufig kann er noch werden?. *Diabetes-J. Schulungsprofi*, Heft 3, 45–47.
- MARKS, H. H., KRALL, L. P. (1971): Onset, course, prognosis and mortality in diabetes mellitus. In: *Joslin's diabetes mellitus*, Marble, A., White, P., Bradley, R. F., Krall, L. P. (Eds.). Philadelphia: Lea & Feabiger, 11. Aufl., 209–254.
- MICHAELIS, D., JUTZI, E. (1991): Epidemiologie des Diabetes mellitus in der Bevölkerung der ehemaligen DDR: Alters- und geschlechtsspezifische Inzidenz- und Prävalenztrends im Zeitraum 1960–1987. *Klin. Med.* **46**, 59–64.
- PANZRAM, G., ZABEL-LANGHENNIG, R. (1981): Prognosis of diabetes mellitus in a geographically defined population. *Diabetologia* **20**, 587–591.
- RUWAARD, D., HOOGENVEEN, R. T., VERKLEIJ, H., KROMHOUT, D., CASPARIE, A. F., VAN DER VEEN, E.A. (1993): Forecasting the number of diabetic patients in The Netherlands in 2005. *Am. J. Public Health* **3**, 989–995.
- SCHNEIDER, H., LISCHINSKI, M., JUTZI, E. (1994): Die Diabetikerprognose im Nordosten Deutschlands. *Z. Ärztl. Fortbild.* **88**, 925–930.
- Statistisches Jahrbuch 1990 der Deutschen Demokratischen Republik (1990), Staatliche Zentralverwaltung für Statistik (Ed.). Berlin: Staatsverlag.
- Statistisches Jahrbuch 1994 für die Bundesrepublik Deutschland (1994). Statistisches Bundesamt (Ed.). Stuttgart: Metzler-Poeschel.
- THOELKE, H., MEUSEL, K. (1994): Zur Überlebensdauer von Typ-II-Diabetikern. Ergebnisse einer 20jährigen Follow-up-Studie in Berlin-Ost. In: *Diskussionsbeiträge zur Gesundheits- und Sozialforschung, Senatsverwaltung für Gesundheit Berlin* (Ed.). Diskussionspapier 19, Berlin.
- ZIMNET, P. (1982): Type2 (non-insulin-dependent) diabetes – An epidemiological overview. *Diabetologia* **22**, 399–411.

## Anhang

Verfahren zum Abschätzen der geschlechtsspezifischen Verteilung der Diabetiker in Abhängigkeit vom Manifestationsalter und der Diabetesdauer:

Wegen der Abhängigkeit der Verweildauer im System vom Manifestationsalter und der Diabetesdauer ist es notwendig, den Anfangsbestand entsprechend aufzugliedern. Dazu wird die Annahme gemacht, daß bei fehlenden Angaben für die Neuerkrankungen in der Vergangenheit diese mit einem konstanten multiplikativen Faktor  $<1$  abnehmen, der altersabhängig anhand der Registerdaten zu bestimmen ist.

1. vorhandene Angaben im Startjahr  $j_0$

$b_{j_0, a, g}$  Bestand nach Alter und Geschlecht

$n_{j_0, a, g}$  Neuerkrankungen nach Alter und Geschlecht

2. Neuzugänge der Vorjahre ( $j < j_0$ )

$$n_{j-1,a,g} = 1.00 \cdot n_{j,a,g}, \quad 0 < \text{Manifestationsalter} \leq 30, \\ 0.97 \cdot n_{j,a,g}, \quad 30 < \text{Manifestationsalter}.$$

Die Neuerkrankungen gehen mit  $j_0 - j$  als Diabetesdauer und  $a$  als Manifestationsalter in den Bestand über

$$b_{j,ma,dd,g} = n_{j,a,j_0-j,g}.$$

3. Berücksichtigung des Sterbegeschehens ( $j < j_0$ )

$$b_{j,ma,dd,g} = b_{j,ma,dd,g} \cdot y(100 - bz_{j,ma,g} + j_0 - j)/y(100 - bz_{j,ma,g})$$

4. Normierung auf den Anfangsbestand ( $j < j_0$ )

$$b_{j,ma,dd,g} = b_{j,ma,dd,g} \cdot b_{j_0,a,g} / \sum_j b_{j,a,g} \quad \text{mit} \quad a = ma + dd.$$

Die Berücksichtigung des Sterbegeschehens erfolgt durch Multiplikation der Bestände mit den zugehörigen Abfallraten ihrer Überlebenskurven, wobei analog zu 2.4 der Bezugspunkt  $bz$  in der Überlebenskurve wegen der Abhängigkeit von der Zeit anzupassen ist. Die anschließende Normierung liefert die Struktur des Bestandes hinsichtlich Manifestationsalter, Diabetesdauer und Geschlecht zu Beginn der Vorhersage.

Eingang des Manuskripts: 18. 3. 98

Anschrift des Verfassers: Peter Heinke, Institut für Diabetes „Gerhardt Katsch“ Karlsburg e.V., Greifswalder Str. 11e, 17495 Karlsburg

# The Reliability Evaluation of Nuclear Imaging Data: An overview on classical ANOVA based approaches

Frank Krummenauer<sup>1)</sup>, Gregor Förster<sup>2)</sup>

## Summary

*Classical ANOVA based parametric methods on the analysis of reliability studies are reviewed and referred to the evaluation of nuclear imaging data. Four important study designs are reviewed and design-adequate estimators for the imaging method's reliability are derived. The methods for estimation and the interpretation within the respective study designs are illustrated by the analysis of recent data arising from a reliability study on the performance of SPECT based somatostatin receptor scintigraphy in Graves' ophthalmopathy.*

## Key words

*Replicate measurements, variance components, item-specific bias, reliability*

## 1 Introduction

Replicate measurements on the same subject can substantially improve the quality of statistical inference intending to evaluate the performance of new imaging methods concerning reproducibility and thus trustability of measurements. Replications will enable investigators to distinguish between different sources of variation and error, such as population variability, item-specific error (that is variation due to interaction between subjects and the diagnostic measurement instruments or observers of interest) and instrumental ("residual") error. Estimation and comparison of the several corresponding variance components can be performed by variance component modelling to be briefly reviewed in the following with respect to several replication designs.

Note, however, that replication can either be used to determine the influence of an instrument's precision or the individual influence of clinical observers on an observation on single subjects. There are various designs of replication studies as for example reviewed by DUNN (1992a, 1992b). Taking four replicate measurements on each study subject can, for example, be regarded as a simple replication study design, whereas the same study could also consist in first taking two replications, then having the study

<sup>1)</sup> Department of Medical Statistics and Documentation, University of Mainz, Obere Zahlbacher Straße 69, D-55131 Mainz, Germany

<sup>2)</sup> Department of Nuclear Medicine, University of Mainz, Langenbeckstraße 1, D-55101 Mainz, Germany



subjects walk around and thus maybe slightly change their physiological conditioning for measurement, and then taking the remaining two replicates afterwards. An extension of this design could next involve the introduction of several clinical observers, who are all asked to take four replicate assessments on each of the study subjects. Next, there would occur a completely different interpretation of replications, if we take four replicate measurements on each subject, but two of them are performed by one observer and the remaining two by another one, who does not know the first observer's results. Further there are staggered designs possible, where not all observers have to see each of the subjects (see e.g. DUNN, 1992a, 1992b or FLEISS, 1981).

These simple examples, which represent quite realistic designs for planning a method performance and feasibility trial, already suffice to indicate, that the kind and interpretation of replications will substantially influence the results of reliability studies and are thus of considerable interest during the planning of corresponding trials.

Without formal specification most measurement models for reliability studies can be regarded as decompositions of type  $X = T + E$ , where  $X$  represents an observer's "observable score" (which is in fact the measured value) and  $T$  the corresponding "true score" of a measurement, which is confounded with a measurement error component  $E$  due to the intrinsic variation of a clinical observer or diagnostic instrument. In this informal notation the reliability can be defined as

$$R = \frac{\text{variance}(T)}{\text{variance}(T + E)} = \frac{\text{variance}(T)}{\text{variance}(T) + \text{variance}(E)},$$

where the latter equality is due to the common assumption of zero correlation between true scores and corresponding measurement errors. Maximum reliability is thus established for  $R = 1$ ; its lower bound  $R = 0$  indicates, that data variation is completely due to measurement error; thus a single assessment would be totally uninformative for the corresponding subject's true score's value. Note that reliability is an important characteristic of clinical measurement procedures. The following sections will thus briefly review approaches for the statistical modelling of reliability and the resulting point estimates (see e.g. LANDIS & KOCH (1975) or FLEISS (1986) for surveys on this field). For an overview on interval estimation in reliability studies the reader is further referred to SHROUT & FLEISS (1979).

Note that the above simplification of a reliability definition also indicates its representation as a formal intraclass correlation, that is the correlation of two measurement replicates on the same subject: Let  $X = T + E$  and  $Y = T + F$  denote corresponding replicates on the same subject with true score  $T$  and independent identically distributed measurements errors  $E$  and  $F$  of the respective assessments. Then the Pearson correlation of  $X$  and  $Y$  becomes

$$\frac{\text{covariance}(T, T)}{\sqrt{\text{variance}(T + E) \cdot \text{variance}(T + F)}} = \frac{\text{variance}(T)}{\sqrt{(\text{variance}(T) + \text{variance}(E))^2}} \equiv R.$$

Corresponding identifications of reliability and intraclass correlation were, for example, exploited in BARTKO (1966) or SHROUT & FLEISS (1979).

## 2 The Simple Replication Design

This initial model is quite simple as compared to its possible extensions already mentioned in the introduction. It is useful to describe data resulting from the so-called

“simple replication trial” (FLEISS, 1986). This study design often arises from an instrumental replication trial, that is replication on each study subject is generated by simply repeating an instrumental measurement or one single clinical observer’s reading on each subject for several subsequent times. One then has to compare two sources of variability, one resulting from within-subject variation (which is in fact due to replication) and the other originated by between-subject variation (which is due to the choice of subjects). A suitable analysis of variance (ANOVA) model for either cases is

$$x_{ij} = s_j + e_{ij} \quad \begin{cases} i = 1, \dots, k \\ j = 1, \dots, n \end{cases}$$

where  $x_{ij}$  is the  $j$ th subject’s realization for replication or observer  $i \in \{1, \dots, k\}$ . The subjects’ main effects  $s_1, \dots, s_n$  are assumed independent and identically normally distributed (i.i.d.) with overall population mean  $\mu$  and variance  $\sigma_s^2$ ; we will write  $s_1, \dots, s_n \sim \mathcal{N}(\mu, \sigma_s^2)$ . The further components  $e_{ij}$  are residual measurement errors, which are *all* assumed i.i.d.  $\mathcal{N}(0, \sigma_e^2)$  and mutually independent from the subjects’ main effects.

The above decomposition of a measurement result  $x_{ij}$  implies, that the observed value  $x_{ij}$  for study subject  $j$ , when obtained in replication  $i$ , is a superposition of the patient’s characteristics ( $s_j$ ), and additionally a residual measurement error  $e_{ij}$  due to the specific configuration when measuring subject  $j$  by replication  $i$ . The total variance of the observed values  $x_{ij}$  therefore becomes  $\sigma_s^2 + \sigma_e^2$ . The basic idea of the classical ANOVA based procedures for reliability evaluation is to exploit the above separation of subject and measurement variation ( $\sigma_s^2$  versus  $\sigma_e^2$ ) to obtain measures for the direct comparison of subject and instrumental or observer variation in terms of intraclass correlations as reliability estimates.

Following the informal introductory definition the (intra–subject) intra class correlation of the above model becomes

$$R = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}$$

and is a substitute to the corresponding reliability  $R$  (FLEISS, 1986). ANOVA decomposition first provides estimates  $\hat{\sigma}_s^2$  and  $\hat{\sigma}_e^2$  for the variance components  $\sigma_s^2$  and  $\sigma_e^2$  by computing the “mean squares”  $MS_s$  and  $MS_e$  for subjects and residual errors, respectively. These can be obtained by any standard software package. The variance component estimates are then

$$\begin{aligned} \hat{\sigma}_s^2 &= \frac{MS_s - MS_e}{k} , \\ \hat{\sigma}_e^2 &= MS_e . \end{aligned}$$

Once having computed the above mean squares, the corresponding variance components are available and can be imputed into the formula for  $R$  to provide a consistent (although slightly biased) reliability estimate

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_e^2} = \frac{MS_s - MS_e}{MS_s + (k - 1) \cdot MS_e} .$$

### 3 The Interexaminer Design

A more attractive study aim might be the comparison of observers instead of only estimating the fraction of loss in reliability due to instrumental or observer variation. Therefore we introduce explicit observer effects into the above ANOVA model. However, one

has to distinguish between two quite different natures of observer effects. The almost natural way would be to regard these effects random, that is the observers under consideration only represent a sample of observers from a much larger pool of possible diagnosticians. The alternative to this random interpretation would be the introduction of fixed observer effects, where the latter call for the interpretation of established professional clinicians as observers, who will remain the only diagnosticians of practical interest for the diagnostic consulting under consideration. Whereas this is a rather unusual assumption in the context of observer reliability, it appears quite attractive in an instrumental comparison scenario: Once having bought, for example, two SPECT (single photon emission computed tomography) instruments for nuclear diagnostics, these instruments will be the only one of practical interest for a quite large population of patients, who will have to undergo SPECT diagnostics. This interpretation is rather that of fixed instrumental bias effects than of random ones. A corresponding multi observer design would consist in the simultaneous consideration of several clinicians having remarkably different diagnostical experience, that is for example the evaluation of a student of medicine's diagnostic readings versus the (consensus) results of several senior pathologists. Whereas the usual theory of reliability estimation rather concentrates on random observer models, this text will therefore also briefly focus on fixed effect estimators.

A suitable ANOVA model for the random case is

$$x_{ij} = s_j + d_i + e_{ij} \quad \begin{cases} i = 1, \dots, k \\ j = 1, \dots, n \end{cases},$$

where  $x_{ij}$  is the  $j$ th subject's realization for replication or observer  $i \in \{1, \dots, k\}$ ; again  $s_1, \dots, s_n \sim \mathcal{N}(\mu, \sigma_s^2)$  refer to the main effects, the  $e_{ij}$  are residual measurement errors assumed i.i.d.  $\mathcal{N}(0, \sigma_e^2)$  and the observer effects  $d_i$  are modeled i.i.d.  $\mathcal{N}(0, \sigma_d^2)$  and mutually independent from the main and residual error effects. The latter means, that we regard different observers as an additional source of variation and thus introduce them as an explaining factor into the overall observable superposition term  $x_{ij}$ .

The corresponding variance components can again be estimated via ANOVA decomposition (LANDIS & KOCH, 1975). The latter results in

$$\begin{aligned} \hat{\sigma}_s^2 &= \frac{MS_s - MS_e}{k}, \\ \hat{\sigma}_d^2 &= \frac{MS_d - MS_e}{n}, \\ \hat{\sigma}_e^2 &= MS_e, \end{aligned}$$

where  $MS_s$ ,  $MS_d$  and  $MS_e$  denote the mean squares for subjects, observers and residual errors, respectively. In this decomposition  $\hat{\sigma}_d^2$  estimates the order of variation, which is due to differences in the mean ratings among observers (it thus estimates "inter-observer bias"), whereas  $\hat{\sigma}_s^2$  estimates the order of variation due to differences among the subjects' mean response over all observers (it thus estimates "inter-subject bias"). The intra class correlation of this model can be estimated consistently, but with bias, by

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_e^2}.$$

Next we regard the observer effects as fixed constants, that is the statistical model reduces to

$$x_{ij} = s_j + \delta_i + e_{ij} \quad \text{with} \quad \sum_{i=1}^k \delta_i = 0 \quad \begin{cases} i = 1, \dots, k \\ j = 1, \dots, n \end{cases},$$

where the non-random observer effects  $\delta_i$  are real numbers, which can be regarded in the same way as fixed constant instrumental biases. SHROUT & FLEISS (1979) provided the corresponding intraclass correlation and a consistent (although biased) estimate. Note that in the fixed observer scenario it is legitimate to isolate some of the raters and compare them to the others by corresponding significance tests, since the raters no longer represent randomly chosen diagnosticians out of a much larger pool of observers, but now are identifiable specialists. A test for absence of observer bias now reduces to the parametric null hypothesis  $H_0 : \delta_1 = \dots = \delta_k = 0$  and can be performed by comparing the observers' variation to the overall residual error, that is the null hypothesis  $H_0$  of absence of inter-rater bias can be tested by referring

$$F = \frac{MS_d}{MS_e}$$

to an F distribution with  $n - 1$  and  $(n - 1)(k - 1)$  degrees of freedom (FLEISS, 1986). If additionally one isolated (maybe less trained or experienced) rater  $i \in \{1, \dots, k\}$  has to be compared to the remaining observers' consensus reading, one may use the linear contrast test statistics

$$C_i = \frac{\bar{X}_{i\bullet} - \frac{1}{k-1} \cdot \sum_{l \neq i} \bar{X}_{l\bullet}}{\sqrt{\frac{k \cdot MS_e}{n(k-1)}}}$$

and, at least for larger sample sizes, establish a significant difference between observer  $i$  and the remaining others as soon as  $|C_i| \geq u_{1-\frac{\alpha}{2}}$ . The latter denotes the  $(1 - \frac{\alpha}{2})$  quantile of the standard normal distribution  $\mathcal{N}(0, 1)$ ; an exact version of this test can be based on Student t distributions, accordingly. As usual a  $\bullet$  subscript indicates summation over the corresponding index, that is, for example,

$$\bar{X}_{i\bullet} = \frac{1}{n} \cdot (X_{i1} + \dots + X_{in}) .$$

#### 4 Item-Specific Bias

A further generalization of the previous interexaminer reliability modelling is the introduction of item-specific bias into the ANOVA decomposition, which has to be estimated by taking replicate measurements on the same subject. In the following we will use the terms *item-specific bias*, *observer/subject interaction* and *instrument/target interaction* interchangeably, since they all refer to the same phenomenon of an observer's deviation from his usual rating patterns when performing replications on a certain subject.

If  $k$  random observers are advised to take  $r < k$  independent replicate readings from each of the study subjects, an appropriate ANOVA model is

$$x_{ijh} = s_j + d_i + (sd)_{ij} + e_{ijh} \quad \begin{cases} i = 1, \dots, k \\ j = 1, \dots, n \\ h = 1, \dots, r \end{cases}$$

where  $x_{ijh}$  denotes the  $j$ th subject's realisation for replication  $h$  by observer  $i \in \{1, \dots, k\}$ . Again  $s_1, \dots, s_n \sim \mathcal{N}(\mu, \sigma_s^2)$  refer to the main effects, the  $e_{ij}$  are i.i.d.  $\mathcal{N}(0, \sigma_e^2)$  residual errors and the observer effects  $d_i$  are i.i.d.  $\mathcal{N}(0, \sigma_d^2)$ . Finally, the random interaction effects  $(sd)_{ij}$  are assumed i.i.d.  $\mathcal{N}(0, \sigma_{sd}^2)$  and mutually independent from all other effects.  $(sd)_{ij}$  represents the influence of item-specific bias, when observer

$i$  is asked to take replicate readings on subject  $j$ . The corresponding variance components can be estimated via ANOVA decomposition, and we obtain the estimates

$$\begin{aligned}\hat{\sigma}_s^2 &= \frac{MS_s - MS_{sd}}{k \cdot r}, \\ \hat{\sigma}_d^2 &= \frac{MS_d - MS_{sd}}{n \cdot r}, \\ \hat{\sigma}_{sd}^2 &= \frac{MS_{sd} - MS_e}{r}, \\ \hat{\sigma}_e^2 &= MS_e,\end{aligned}$$

where  $MS_{sd}$  denotes the mean squares for the subject/observer interaction. The intra class correlation can now be estimated consistently, but with bias, as

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_{sd}^2 + \hat{\sigma}_e^2}.$$

It should be mentioned, that, from a formalistic point of view, *any* observer reliability study dealing with more than one observer seeing the study subjects should be *modeled* in the above way, since observer/subject interaction can never be excluded in advance. Seperate estimation of these effects, however, is not possible unless replicate measurements of the observers are available (e.g. SHROUT & FLEISS, 1979). Since sometimes replications by each observer is too cost-extensive or even impossible due to ethical restrictions, we presented the previous expedient solutions in section 3 disregarding the interaction effects  $(sd)_{ij}$ . If  $R$  is estimated without correction for item-specific bias, then item-specific bias and residual error will be confounded and the corresponding variance component estimate  $\hat{\sigma}_e^2$  might turn out remarkably increased as compared to its item-specific bias corrected analogue.

## 5 An Example: The SPECT Data

The most important of the above models will be illustrated within the analysis of a recent reliability study on single photon emission computed tomography (SPECT) of somatostatin receptor scintigraphy (SCR) in patients with Graves' ophthalmopathy (FÖRSTER et al., 1998). The SPECT-based SCR allows for the assessment of orbital inflammation of patients with thyroid eye disease. For SRC the radio-labelled somatostatin analogue  $^{111}\text{[In-DTPA-D-Phe]}$  octreotide is used most frequently. The development of this synthetic analogue with a much longer lifetime than somatostatin enables to demonstrate SR positive tissue and activated T cell accumulation. Patients with active thyroid eye disease showed remarkably increased orbital accumulation of octreotide. To properly quantify the orbital uptake, however, SPECT images are necessary. On the other hand, the common disadvantage of orbital SPECT data is the absence of exact anatomical proportions of the orbitae. In this trial we therefore focussed on the evaluation of reliability data in the SPECT-based SCR setting.

Indium-111-pentetreotide scintigraphy was performed on 8 consecutive patients (four female, four male, median age 55 years) with active thyroid eye disease. All patients had an untreated ophthalmopathy of recent onset (less than six months). SPECT images were obtained with a large field-of-view gamma camera four hours after intravenous injection of 110 MBq (3 mCi)  $^{111}\text{[In-DTPA-D-Phe]}$  pentetreotide with an effective dose equivalent of 13 mSv. The gamma camera was equipped with a medium-energy parallel-

hole collimator. SPECT was performed in a stop-and-shoot mode. 120 30-sec-views per head were acquired over a  $360^\circ$  rotation in a  $128 \times 128$  pixel matrix. The SPECT data was then processed using filtered backprojection and attenuation correction was done with a coefficient of 0.9/cm (Chang algorithm). Transversal SPECT slices parallel to the middle of the orbita were reconstructed resulting in two-pixel slice thickness with a voxel size of  $4.67 \times 4.67 \times 9.34$  mm. An optimal orbital image with most activity within the region of the orbita was selected and predetermined irregular regions of interest (ROIs) for both orbitae were positioned, yielding an obvious origin of variability due to observer experience, item-specific bias and thus the necessity of taking replicate measurements on each patient. Therefore the SPECT data of 8 patients was evaluated by four independent observers 3 to 4 times each. However, to keep the following computations more instructive, we omit the imbalanced fourth replication of one observer and thus rely on three replicates on each patient by each of the four observers. The result of each single measurement was expressed in counts per voxel [cts/vox]. For further details on the generation of the SPECT data see FÖRSTER et al. (1998).

To illustrate the performance of the models in sections 2, 3 and 4 we consider the following three steps in analysis. First the reliability is computed for the obviously most correct model incorporating correction for item-specific bias; note that subject/observer interaction at least has to be expected due to the somewhat orbitae dependent feasibility of ROI positioning during replication. Therefore we estimate the SPECT-based SCR reliability for  $r = 3$  replications of  $k = 4$  observers on  $n = 8$  patients out of the following model:

$$x_{ijh} = s_j + d_i + (sd)_{ij} + e_{ijh} \quad \begin{cases} i = 1, \dots, 4 \\ j = 1, \dots, 8 \\ h = 1, \dots, 3 \end{cases}.$$

First we compute the variance components as postulated in section 4, where the mean standard errors can be easily generated by using PROC GLM of the standard statistical software SAS. Based on the SAS output results we obtain the variance component estimates

$$\hat{\sigma}_s^2 = \frac{MS_d - MS_{sd}}{k \cdot r} \approx \frac{20190 - 348}{4 \cdot 3} \approx 1654,$$

$$\hat{\sigma}_d^2 = \frac{MS_s - MS_{sd}}{n \cdot r} \approx \frac{4664 - 348}{8 \cdot 3} \approx 180,$$

$$\hat{\sigma}_{sd}^2 = \frac{MS_{sd} - MS_e}{r} \approx \frac{348 - 148}{3} \approx 67,$$

$$\hat{\sigma}_e^2 = MS_e \approx 148$$

and immediately recognize, that most variation contained in the data at hand is due to subject inhomogeneity. Item-specific bias is of quite ignorable order explaining only 3% of the total variance as recognized by

$$\theta = \frac{\hat{\sigma}_{sd}^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_{sd}^2 + \hat{\sigma}_e^2} \approx \frac{67}{1654 + 180 + 67 + 148} \approx 0.03.$$

The overall intraclass reliability estimate becomes

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_{sd}^2 + \hat{\sigma}_e^2} \approx \frac{1654}{1654 + 180 + 67 + 148} \approx 0.81$$



indicating rather moderate reliability of the SPECT-based SCR. The main cause is the somewhat large variance component due to observer variation as estimated by

$$\frac{\hat{\sigma}_d^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_{sd}^2 + \hat{\sigma}_e^2} \approx \frac{180}{1654 + 180 + 67 + 148} \approx 0.09 .$$

Summarizing we have quite small item-specific bias, but non-ignorable observer variation. Therefore SPECT based SCR should only be applied by rather experienced physicians or especially well-trained assistants. These professionals, however, will provide remarkably reliable results ( $\theta = 3\%$ ) as will also be motivated in the subsequent analyses.

To illustrate this issue we will go on to analyse the above data in a more exploratory way: First we aggregate the observers' replications on one subject by computing mean values, that is we artificially reduce the study design to the model of section 3. If we further assume random observer effects (that is we do not regard any of the observers as an isolated specialist in SPECT evaluation), then the variance components based on the SAS procedure PROC GLM turn out as follows:

$$\begin{aligned} \hat{\sigma}_s^2 &= \frac{MS_d - MS_e}{n} \approx \frac{6730 - 116}{4} \approx 1654, \\ \hat{\sigma}_d^2 &= \frac{MS_s - MS_e}{k} \approx \frac{1555 - 116}{8} \approx 180, \\ \hat{\sigma}_e^2 &= MS_e \approx 116 . \end{aligned}$$

Note that we cannot estimate the item-specific bias component, which is formally confounded with the residual error component due to the artificial ignorance of the original replication design. On the other hand, the residual error component turns out even much smaller than in the previous model, which can be explained by the overly artificial reduction in variation due to averaging the observers' replicate readings. Therefore the overall reliability point estimate must even be expected to increase as compared to the previous scenario. Accordingly, we have an obviously large observer related variance component  $\hat{\sigma}_d^2 = 180$ , which results in the overall reliability estimate of

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_d^2 + \hat{\sigma}_e^2} \approx \frac{1654}{1654 + 180 + 116} \approx 0.85 .$$

Again we observe rather moderate reliability of the SPECT evaluation mainly caused by observer variation. Next one has to ask, whether this variation is clinically relevant. Note that the above observer variability does not necessarily indicate statistically significant differences between the observers under consideration! One has to distinguish between the mean measurement, which can be increased or decreased by one or more of the observers as compared to the remaining ones (deviation in accuracy) and possibly moderate reliability, which is rather a fact of deviations in precision.

To obtain information on observer differences and to become able to identify a possibly incompatible observer by contrasting his results to those of the remaining ones, we switch to the fixed observers model in section 3 and take a short descriptive look on the four identifiable observers' mean ratings over the 8 patients under consideration.

Table 1 shows these mean values and the corresponding standard deviations; it illustrates the obvious increase in mean measurements contained in the data provided by observer 4 as compared to the other diagnosticians' results. Therefore we use the contrast statistic  $C_4$  of section 3 to assess the statistical significance of this deviation. We

Table 1: minima, maxima, means and standard deviations for observers 1–4 over  $n = 8$  patients, where each single value resulted from three replicate readings on each patient by the respective observer

observer	minimum	maximum	mean	standard dev.
1	58	208	137	44
2	74	186	144	37
3	63	187	144	41
4	86	214	170	45

obtain

$$\begin{aligned}
 C_4 &= \frac{\bar{X}_{4\bullet} - \frac{1}{k-1} \cdot (\bar{X}_{1\bullet} + \bar{X}_{2\bullet} + \bar{X}_{3\bullet})}{\sqrt{\frac{k \cdot MS_e}{n(k-1)}}} \\
 &\approx \frac{170 - \frac{1}{3} \cdot (139 + 144 + 144)}{\sqrt{\frac{4 \cdot 116}{8 \cdot 3}}} \\
 &\approx 4.4,
 \end{aligned}$$

which dramatically exceeds the standard normal quantile  $u_{0.995} = 2.326$  as well as the corresponding  $t$  distribution's quantile for  $\alpha = 1\%$ . Therefore observer 4 significantly differs from the other observers' collection at the 1% significance level. Note, however, that one cannot necessarily conclude, that observer 4 is the least experienced one among the four observers! Maybe observer 4 is the most competent one among all diagnosticians under consideration, and his ROIs are optimum valid providing maximum large counts. Nevertheless, his rather large standard deviation as shown in Table 1 rather contradicts superiority of observer 4. In general, one should rather contrast all 4 observers' ratings to the results of a possible gold standard to obtain corresponding information on the validity of their results.

As a final exercise we illustrate the simple replication reliability formula in section 2 by completely perverting the original study design. We compute the "intra-observer" reliability for each rater, respectively, by separately considering his three replications on each of the eight subjects and thus stratifying the primary reliability analysis by observers. Note that the intraclass coefficient in section 2 now becomes

$$\hat{R} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_e^2} = \frac{MS_s - MS_e}{MS_s + (r-1) \cdot MS_e},$$

since replication on subjects is now due to the  $r = 3$  replicate readings by each observer. Computing the mean squares with PROC GLM now provides reliability estimates of 87%, 88%, 97% and 96% for observers 1, 2, 3 and 4, respectively. Surprisingly, this design shows the quite satisfying "intra-observer reliability" of observer 4 as compared to observers 1 and 2. Note, however, that the latter arguments hardly represent a sensible analysis of the primary study design; they rather provide an additional insight into the data and are merely presented for the sake of illustration. Despite deviation from the primary design there is another pitfall contained in the previous intra-observer computation as already mentioned by FLEISS (1986): If replicate measurements on the same subject are taken by the same observer, the above reliability estimate cannot distinguish

between a really reliable assessment and an implicit or even explicit influence of rater memory. A well-trained nuclear physician will be able to recognize patients by only seeing their orbital image and may thus be seduced to imitate his previous diagnostic environment (e.g. during re-positioning the ROI); this is in fact not intended by the assessment of reliability. Therefore study designs and measurement circumstances should always be triggered in a way, which optimally avoids such effects.

## 6 Discussion

This text tried to survey classical elementary methods for the assessment of reliability in clinical and, in particular, nuclear imaging data by providing elementary formulae for the computation of reliability estimates. These formulae can be easily applied on the basis of software outputs as generated by PROC GLM in SAS or corresponding modules in other software packages. Therefore this text rather tried to clarify decisions on the choice of an appropriate model (and thus a design-adequate reliability estimate) than to provide extensive and previously published formulae, which can be simply avoided by appropriate application of standard statistical software. In particular the lengthy analysis of the SPECT data in section 5 intended to rather explain the different directions in *interpretation* of the several reliability computations than to repeat the well-known ways of *computations*. Due to the same reasons we omitted the repetition of corresponding formulae for confidence intervals, which have been summarized by SHROUT & FLEISS (1979). SAS macros for the design-adequate computation of point and interval reliability estimates are available from the authors on request.

Of course, such an overview cannot be complete. It should rather provide simple dictionary suggestions for approaches to reliability estimation in standard study designs involving replicate measurements. Technometrical literature provides corresponding alternatives for designs, which exceed the restrictions introduced here, but this overview focussed on the most frequent study designs to keep the ideas of reliability estimation more instructive. For example, we did not provide extensions for the estimation of reliability, when data happens to be incomplete; the reader is referred to the informative works of THOMPSON & ANDERSON (1975) or FLEISS (1981), which illustrate the use of ANOVA decomposition in the context of missing or unbalanced data, such as e.g. intended in the balanced incomplete block designs of FLEISS (1981). Next we only considered very special cases of the quite general mixed model theory, where e.g. SEARLE (1987) provides a survey on estimation methods for these models. Nevertheless, much practical data applications will on the contrary allow or even call for very simple analysis due to the possibly uncontrollable sources of bias and errors inevitable in clinical reality.

Nevertheless, the assumptions contained in this paper are restrictive. Note, for example, that the standard ANOVA models stressed in sections 2–4 assume observers to produce (multiple) readings without any relative bias. This may be a sensible assumption in the setting of the simple replication design, but is an obviously hard and thus questionable restriction, when different (human) examiners are being dealt with. On the other hand, this assumption allows for the use of the intraclass *correlation*  $\rho$  as the main endpoint in the comparison of multiple readings on one subject: The usual important critics in the application of correlation coefficients in the comparison of intra-individually dependent data series (such as resulting from method comparison studies) are now simply weakened by modeling and thus postulating their absence. Nevertheless, one should always remember this issue and therefore carefully consider, whether the assumptions implicitly and explicitly contained in ANOVA approaches can be regarded justified in the data at hand. Next any of the very simple models mentioned here assumes (multivariate) normally distributed data, i.e. this overview does not consider the problem of robustness

and, for example, nonparametric approaches to the analysis of multi observer trials. Literature provides a lot of solutions to this issues. This review, however, rather intended (as already mentioned) to illustrate some principle issues in the analysis of reliability studies, which are of practical relevance for the non-statistical directions in clinical medicine, too, such as clinical imaging or quantitative pathology.

Further it should be mentioned, that only one choice of methods for analysis may not be appropriate in practice; one should rather try to obtain a sufficient descriptive overview on the data at hand and then try to fit appropriate variance (or covariance) component models, such as illustrated in section 5 or, for example, in the confirmatory factor analysis described by LONG (1983a, 1983b).

Another possible source of confusion, which is hardly to control for, is the population dependence of reliability: Nuclear imaging methods may be highly efficient and reliable in patients with severe orbital dysfunction, but undergo dramatic loss in reliability as soon as being applied to patients with moderate or different sources of ophthalmological disorder. Therefore planning of reliability studies should sensitively incorporate the main goal of the study intention, that is either providing a representative range of observers and instruments and a maximum homogeneous study population (if measurement errors are focussed on and item-specific bias has to be controlled for) or providing a representative range of study subjects to allow for estimation of subject-specific properties, too. These study designs correspond to diagnostic score evaluation studies (which primarily intend to assess rater variation as a measure of reliability of the proposed diagnostic score) and diagnostic studies (which are primarily interested in a subject's optimum reliable diagnostic findings and adjust the latter for rater variation).

In summary, results of reliability studies are not only depending on the data at hand, but are also influenced by both the study design and the underlying population chosen in advance, and last (but certainly not least as indicated in section 5) on the clinical context of actual interest.

The relation of reliability to quality management and assurance is obvious, in particular for clinical disciplines, which are based on imaging diagnostics (such as nuclear medicine, radiology, cardiology, orthodontics or neurosurgery). Therefore reliability studies can be regarded as an important, although much underused, tool for the evaluation of diagnostic agents or instruments in clinical medicine. This review tried to provide an overview on classical ANOVA based approaches; however, a lot of subsequent statistical effort will be necessary and desirable to make this field become an established and common tool in the context of diagnostic trials.

### Acknowledgement

The authors are grateful to Prof. Dr. JÖRG MICHAELIS (Institut für Medizinische Statistik und Dokumentation der Universität Mainz), Prof. Dr. MARKUS LÖFFLER (Institut für Medizinische Informatik, Statistik und Epidemiologie der Universität Leipzig) and Prof. Dr. STEFAN WELLEK (Zentralinstitut für seelische Gesundheit Mannheim) for numerous constructive suggestions on the preparation of this manuscript. We further thank Prof. Dr. GEORGE KAHALY (Klinik und Poliklinik für Endokrinologie und Stoffwechselkrankheiten der Universität Mainz) for his kind help in recruitment of patients suffering from Graves' ophthalmopathy.

### References

- BARTKO, J. J. (1996): The intraclass correlation coefficient as a measure of reliability. *Psychological Reports* **19**, 3–11

- DUNN, G. (1992a): Design and analysis of reliability studies. *Statistical Methods in Medical Research* **1**, 123–157
- DUNN, G. (1992b): *Design and Analysis of Reliability Studies*. Edward Arnold, London
- FLEISS, J. L. (1981): Balanced incomplete block designs for inter-rater reliability studies. *Applied Psychological Measurement* **5**, 105–112
- FLEISS, J. L. (1986): *The Design and Analysis of Clinical Experiments*. John Wiley & Sons, New York
- FÖRSTER, G., KRUMMENAUER, F., HABEREM, G., ANDREAS, J. and KAHALY, G. (1998): Inter- and Intra-Observer Variability of Somatostatin-Receptor Scintigraphy in Thyroid Eye Disease (TED). *Thyroid* **8**, 456
- LANDIS, J. R. and KOCH, G. G. (1975): A review of statistical methods on the analysis of data arising from observer reliability studies (parts I and II). *Statistica Neerlandica* **29**, 101–123 and 151–161
- LONG, J. S. (1983a): *Confirmatory Factor Analysis: A Preface to LISREL*. Sage Publications, Beverly Hills and London
- LONG, J. S. (1983b): *Covariance Structure Models: An Introduction to LISREL*. Sage Publications, Beverly Hills and London
- SAS Software: SAS User's Guide (Version 6, 4th edition). SAS Institute Inc., Cary, North Carolina, U.S.A.
- SEARLE, S. R. (1987): *Linear Models for Unbalanced Data*. John Wiley & Sons, New York
- SHROUT, P. E. and FLEISS, J. L. (1979): Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin* **86**, 420–428
- THOMPSON, W. O. and ANDERSON, A. L. (1975): A comparison of designs and estimators for the two-stage nested random model. *Technometrics* **17**, 37–44

Eingang des Manuskripts: 28. 1. 99

Anschrift der Verfasser: Dr. Frank Krummenauer, Institut für Medizinische Statistik und Dokumentation der Universität Mainz (Direktor: Prof. Dr. J. Michaelis), Obere Zahlbacher Straße 69, D-55131 Mainz, Tel.: (061 31) 17 31 07, Fax: (061 31) 17 29 68, email: krummi@imsd.uni-mainz.de. Dr. Gregor Förster, Klinik und Poliklinik für Nuklearmedizin der Universität Mainz (Direktor: Prof. Dr. P. Bartenstein), Langenbeckstraße 1, D-55101 Mainz, Tel.: (061 31) 17 71 24, Fax: (061 31) 17 23 86, email: foerster@nuklear.klinik.uni-mainz.de

## Effekte fehlerhafter Expositions- und Confounderbestimmungen in epidemiologischen Studien

Gerd Ripplin<sup>1)</sup>

### Zusammenfassung

*Innerhalb der Epidemiologie ist es unerlässlich, Studienergebnisse kritisch zu hinterfragen und potentiellen Ergebnisverzerrungen nachzugehen. Ungenaue Expositions- oder Confoundererhebungen können zu verfälschten Resultaten führen. Es wird dargestellt, wie bei der logistischen Regression oder anderen statistischen Verfahren Tests ungültig und Schätzungen systematisch verzerrt werden können.*

*Da bei vielen Expositions- oder Confounderbestimmungen ein Erhebungsfehler angenommen werden muß, ist das Auftreten solcher Effekte eher die Regel als die Ausnahme und Epidemiologen sollten in ihren Publikationen auf sie hinweisen und sie diskutieren. Nach Möglichkeit sollte bereits im Studiendesign eine Fehlerkorrektur vorgesehen und bei der Auswertung berücksichtigt werden.*

### Schlüsselworte

*Meßfehler, Mißklassifikation, Verzerrung, Epidemiologische Studien*

### Summary

*In epidemiology it is essential to assess critically potential biases of study results. In particular, inaccurate exposure or confounder measurements can lead to false findings. It is described, in which way tests and estimates using the logistic regression model or other statistical methods can become invalid or biased respectively. Since for many exposure or confounder measurements an additional error must be stated, the occurrence of such effects is rather the rule than the exception and epidemiologists should discuss these in their publications. If possible, a correction for measurement error should be considered at the design planning stage and eventually in the final analysis.*

### Keywords

*Measurement error — Misclassification — Bias — Epidemiologic studies*

---

<sup>1)</sup> Institut für Medizinische Statistik und Dokumentation der Johannes Gutenberg-Universität Mainz



## Einleitung

Die Epidemiologie hat eine Vielzahl von Methoden entwickelt, um zu quantitativen Schlußfolgerungen über Einflußgrößen für das Auftreten einer Krankheit zu gelangen. Diese Schlußfolgerungen sind jedoch potentiellen Verzerrungen ausgesetzt. Insbesondere können durch ungenaue Expositions- und Confoundererhebungen, oder durch Surrogatmarker systematisch falsche Schlußfolgerungen resultieren. Übliche Analysemethoden wie die logistische Regression oder andere Test- und Schätzmethoden setzen voraus, daß die betrachteten Merkmale fehlerfrei gemessen werden. Es ist jedoch eher die Regel als die Ausnahme, daß diese Voraussetzung nicht erfüllt ist. Z. B. können bei Interviews und Fragebögen aufgrund von Verständnisschwierigkeiten oder aufgrund einer mangelnden Konzentration Fragen ungenau oder falsch beantwortet werden. Ebenso können Erinnerungsschwierigkeiten (bzw. ein Erinnerungszwang!) zu fehlerhaften Angaben führen. Darüber hinaus kann es insbesondere bei „heiklen“ Fragen zu bewußt falsche Angaben kommen (z. B. bei Fragen zum Alkohol- oder sonstigem Drogenkonsum). Auch bei anderen ungenauen Meßinstrumenten wie z. B. physikalischen Meßmethoden muß ein gewisser Meßfehler immer in Kauf genommen werden. Als Beispiele können Messungen der Radonexposition oder von elektromagnetischen Feldern genannt werden. Noch schwieriger ist es, Parameter des Lebensstils wie die körperliche Aktivität (z. B. bei Rheumastudien) oder Ernährungsparameter quantitativ zu erfassen. Das Problem einer exakten Messung wird zusätzlich von der Variabilität der Meßwerte über die Zeit innerhalb der Patienten/Probanden verstärkt (z. B. Blutdruck- oder Cholesterinbestimmung). Zuletzt besteht oft auch eine Unsicherheit in der Bezeichnung Exposition: Welcher Abstand zum Kernkraftwerk soll noch unter dem Begriff „exponiert“ subsumiert werden? Ab welchem Alter ist eine Schwangerschaft als eine „Risikoschwangerschaft“ zu bezeichnen?

Die Effekte ungenauer Expositions- oder Confoundererhebungen sind somit fast allgegenwärtig. Im folgenden wird daher eine kurze Zusammenfassung dieser Effekte gegeben. Um als Wissenschaftler die potentiellen Verzerrungen eigener oder in der Literatur beschriebener Studien besser beurteilen zu können, ist es nötig, genaue Kenntnisse über die Auswirkungen fehlerhafter Expositions- oder Confoundererhebungen zu besitzen.

Wenn eine Variable nur ungenau erfaßt werden kann, dann wird bei stetigen Merkmalen von einem Meßfehler gesprochen, wohingegen bei kategoriellen Daten der Begriff der Mißklassifikation gebräuchlich ist. Als Beispiel können der Raucherstatus, z. B. codiert in „Raucher“ vs. „Nichtraucher“, und eine stetig erfaßte Exposition, z. B. gegenüber einer bestimmten Chemikalie, genannt werden. Im ersten Fall können Mißklassifikationen und im zweiten Fall Meßfehler auftreten. Im folgenden soll jedoch der Begriff Meßfehler auch den kategoriellen Fall miteinschließen.

Die Auswirkungen fehlerhafter Variablen können anhand einer Kontingenztafel der beiden binären Variablen *Raucherstatus* und dem Vorliegen einer *Krankheit* beispielhaft betrachtet werden. Wenn der Raucherstatus extrem mißklassifiziert würde, so daß der wahre Status nur mit einer Wahrscheinlichkeit von 50% erkannt wird, dann ist leicht einzusehen, daß z. B. ein geschätztes Odds-ratio verzerrt bestimmt wird und ein  $\chi^2$ -Test keine Abhängigkeit mehr zwischen beiden Variablen erkennen könnte, da dann rein zufällig die Ausprägung „Raucher“ bzw. „Nichtraucher“ zu beobachten wäre.

Von besonderer Wichtigkeit bei der Beschreibung der Effekte fehlerhafter Messungen ist die Unterscheidung von differentiellen vs. nichtdifferentiellen Fehlern. Eine Fehlerstruktur ist dann nichtdifferentiell, wenn gilt:

$$f_{Z|X,Y}(z|X=x, Y=y) = f_{Z|X}(z|X=x),$$

bzw.

$$P_{Z|X,Y}(z|X=x, Y=y) = P_{Z|X}(z|X=x).$$

Dabei seien  $Y$  und  $X$  Variablen, die fehlerfrei beobachtbar sind, wobei statt der wahren Variablen  $X$  aber nur die Variable  $Z$ , eine fehlerhafte Form von  $X$ , für die Analyse zur Verfügung steht. Obige Formel besagt, daß ein nichtdifferentieller Fehler, der in der Variablen  $Z$  ist, nur abhängig von  $X$  sein darf. Ein nichtdifferentieller Fehler bezüglich  $Y$  ist ein Fehler, der unabhängig von  $Y$  ist. In eine epidemiologische Fragestellung übersetzt, bedeutet ein differentieller Fehler bezüglich des Fall-Kontroll Status ( $Y$ ), daß in Abhängigkeit von diesem Status Fehler mit einer unterschiedlichen Struktur auftreten. Bei einem nichtdifferentiellen Fehler in der Variablen  $X$  hingegen ist die Fehlerstruktur bei Fällen und Kontrollen identisch. Der letztere Fall kann in der Regel bei Kohortenstudien angenommen werden, da die Expositionsbestimmung vor der Kenntnis einer späteren Erkrankung erfolgt und somit unabhängig von ihr sein sollte. Bei Fall-Kontrollstudien kann z. B. bei Messungen von Radonkonzentrationen ein nichtdifferentieller Fehler angenommen werden, da die Genauigkeit der Meßmethode für Fälle und Kontrollen als identisch angenommen werden kann. Differentielle Fehler können z. B. durch Fragebögen oder Interviews hervorgerufen werden, bei denen die sensibilisierten Fälle ihre Expositionsabschätzungen eher aufrunden. Zusätzlich könnten Kontrollen diesen Expositionen womöglich weniger Bedeutung zusprechen und eventuell eher abrunden. Ein solcher Mechanismus systematischer Fehler würde allein aufgrund der Erhebungsfehler die zu untersuchende Exposition als schädlich erscheinen lassen.

In den folgenden Abschnitten werden die Effekte fehlerbehafteter Variablen genauer aufgezeigt. Dieses Problem wurde erstmals von BROSS (1954) aufgegriffen. Er untersucht die Auswirkungen von Mißklassifikationen in einer Vierfeldertafel. Sowohl Schätzungen als auch Tests können beeinträchtigt werden. Auch der Trendtest kann hinsichtlich der Effekte von Mißklassifikationen untersucht werden. Im multivariaten Fall kann gezeigt werden, daß Meßfehler im Confounder mindestens zum partiellen Verlust der Confounderkontrolle führen. Somit können auch Effektschätzungen und Tests bezüglich der Exposition ihre nominalen Eigenschaften verlieren.

### Mißklassifikationen und der $\chi^2$ -Test bei $2 \times 2$ -Tafeln

BROSS (1954) (vgl. auch NEWELL, 1963) untersuchte den Zusammenhang zwischen Mißklassifikationen in Vierfeldertafeln und der Anwendung des  $\chi^2$ -Tests. Man kann sich die von ihm untersuchte Situation als einen Vergleich zweier Binomialverteilungen vorstellen. Dieses mathematische Modell geht davon aus, daß die Gruppenzugehörigkeit (z. B. Fall/Kontrolle) immer ohne Fehler erkannt wird, während der Ausgang des Binomalexperimentes (Expositionsbestimmung) durch eine eventuelle Fehlklassifikation überlagert sein kann.

Die mathematische Formulierung lautet: Sei  $x$  binomialverteilt und gebe die beobachtete Anzahl der „Ereignisse“ wieder (z. B. beobachtete Anzahl exponierter Personen). Sei ferner  $u$  die tatsächliche Anzahl der Ereignisse,  $v$  die Anzahl der Personen, die fälschlicherweise ohne Ereignis gewertet wurden und  $w$  die Anzahl der fälschlicherweise als mit einem Ereignis eingestuft. Daraus folgt:  $x = u - v + w$ .

Für die entsprechenden Wahrscheinlichkeiten gelte:

$$p = P(\text{Ereignis}),$$

$$\theta = P(\text{als ohne Ereignis klassifiziert} \mid \text{Ereignis liegt vor}) = 1\text{-Sensitivität},$$

$$\phi = P(\text{als Ereignis klassifiziert} \mid \text{Ereignis liegt nicht vor}) = 1\text{-Spezifität}.$$

Dann folgt unmittelbar, daß sich der Erwartungswert  $E\left(\frac{x}{n}\right)$  als

$$E\left(\frac{x}{n}\right) = p - p\theta + (1-p)\phi$$

berechnen läßt.

Der Schätzer  $\frac{x}{n}$  ist also i. a. verzerrt für  $p$ , wobei die Richtung des Bias von der jeweiligen Parameterkonstellation abhängig ist. Damit ist auch die naive Varianzschätzung i. a. verzerrt.

Sei nun  $x_1$  die beobachtete Anzahl der Erfolge der Fälle und sei  $x_2$  analog für die Kontrollen definiert. Untersucht man die Differenz  $d = \frac{x_1}{n_1} - \frac{x_2}{n_2}$ , dann hat, wie aus der Formel für  $E\left(\frac{x}{n}\right)$  unmittelbar ersichtlich, unter der Nullhypothese  $p_1 = p_2$  die Differenz noch immer den Erwartungswert Null. Zudem ist auch die Varianzschätzung der Differenz unverzerrt. Bross argumentiert nun, daß der  $\chi^2$ -Test asymptotisch äquivalent zu einem Test basierend auf  $d$  und deshalb gültig ist. Allerdings folgt aus den Mißklassifikationen ein Powerverlust.

Implizit geht Bross von nichtdifferentiellen Fehlern aus. Bei differentiellen Fehlern müßten unterschiedliche Fehlerraten  $\phi_1$  und  $\phi_2$  bzw.  $\theta_1$  und  $\theta_2$  (je nach Gruppenzugehörigkeit) anstatt  $\phi$  und  $\theta$  berücksichtigt werden. Bei differentiellen Fehlern wird das Niveau des  $\chi^2$ -Tests jedoch i. a. nicht eingehalten.

MOTE & ANDERSON (1965) (vgl. auch ASSAKUL & PROCTOR, 1967) verallgemeinerten das von Bross untersuchte Problem auf beliebige  $i \times j$ -Tafeln. Der  $\chi^2$ -Test bleibt auch im allgemeinen  $i \times j$ -Fall unter der Annahme nichtdifferentieller Fehler gültig. Allerdings bleibt auch der Powerverlust bestehen. Bei Mißklassifikationen in nur einer der beiden Variablen geben Mote und Anderson eine asymptotische Formel für den Powerverlust an.

### Mißklassifikationen und Effektschätzungen in 2×2-Tafeln

Neben der von Bross überprüften Gültigkeit des  $\chi^2$ -Tests sind ebenso seine und andere Untersuchungen der durch Meßfehler beeinflussten *Schätzungen*, z. B. des Odds-ratios, von zentralem Interesse.

Während beim  $\chi^2$ -Test (unter der Voraussetzung nichtdifferentieller Fehler) das einigermaßen beruhigende Ergebnis von gültigen Tests vorgewiesen werden kann (wenngleich mit einem gewissen Powerverlust), so ist, was die Schätzung anbetrifft, die Lage weniger erfreulich. Es wäre ein relativ schönes Resultat, wenn hier als Ergebnis nachgewie-

Tabelle 1: Effekte (nicht-)differentieller Fehler bei der Vierfeldertafel

	Wahre Kontingenztafel		Beobachtete Kontingenztafeln			
			Nichtdiff. Fehler		Diff. Fehler	
	Lungenkarzinom		Lungenkarzinom		Lungenkarzinom	
	ja	nein	ja	nein	ja	nein
Raucher	70	30	63	27	70	24
Nichtraucher	30	70	37	73	30	76
$\Sigma$	100	100	100	100	100	100
Odds-ratio	5,44		4,60		7,39	
$\chi^2$ -Teststatistik	32,0		26,0		42,5	

sen werden könnte, daß die Schätzungen unverzerrt sind mit einer vielleicht vergrößerten Varianz, aber in der Realität sind auf fehlerhafte Variablen beruhende Schätzungen zusätzlich zur Varianzvergrößerung auch noch verzerrt: Unter der Voraussetzung nichtdifferentieller Fehler sind die Risikodifferenz, das relative Risiko und das Odds-ratio zur Nullhypothese hin verzerrt (KEYS & KIHLEBERG, 1963, GULLEN et al., 1968). Das gilt auch, wenn *beide* Variablen nichtdifferentielle Fehler aufweisen.

Unter der Annahme differentieller Fehler ist eine Verzerrung nach beiden Richtungen möglich, so daß in diesem Fall keine allgemeine Aussagen über die Richtung des Bias erfolgen kann.

Die obige Tabelle veranschaulicht den Effekt einer nichtdifferentiellen bzw. differentiellen Mißklassifikation des Raucherstatus. Während im ersten Fall Nichtraucher ihren Status fehlerfrei angeben ( $\Rightarrow$  Spezifität =  $Sp = 1$ ), geben Raucher in 10% der Fälle an, nicht zu rauchen ( $\Rightarrow$  Sensitivität =  $Se = 0,9$ ). Daraus folgt eine mißklassifizierte Kontingenztafel, für die das Odds-ratio um ca. 15% erniedrigt ist. Anhand der Reduzierung der Prüfgröße des  $\chi^2$ -Tests ist der Powerverlust ersichtlich. Allerdings fällt dieser bei diesem speziellen Beispiel aufgrund des deutlichen Zusammenhangs zwischen Raucherstatus und dem Auftreten eines Lungenkarzinoms nicht sonderlich ins Gewicht. Wenn angenommen wird, daß die Angaben der sensibilisierten Fälle mit  $Se(\text{Fälle}) = Sp(\text{Fälle}) = 1$  erhoben werden könnten, aber die Kontrollen einen positiven Raucherstatus nur zu 80% auch angeben, also  $Se(\text{Kontr.}) = 0,8$ ,  $Sp(\text{Kontr.}) = 1$ , dann würde der Raucherstatus differentiell mißklassifiziert werden. Es resultiert sowohl ein erhöhtes Odds-ratio als auch ein zu großer Wert der  $\chi^2$ -Teststatistik, so daß der p-Wert zu klein ausfällt.

Sensitivität und Spezifität als Maße der Güte von Meßmethodiken binärer Variablen haben nicht-identische Wirkungen auf den Bias z. B. einer Odds-ratio-Schätzung. Wenn mehr nicht-exponierte als exponierte Personen in einer Studie sind, beeinflußt die Spezifität deutlicher den Bias einer auf der beobachteten Kontingenztafel beruhenden Schätzung, denn selbst bei einer relativ hohen Spezifität von beispielsweise 0,95 würden dann sehr viele Personen fehlklassifiziert werden.

Der Bias von Schätzungen wird neben der Sensitivität und der Spezifität der Meßverfahren auch von der Prävalenz der Exposition bzw. der Krankheit beeinflußt. So zeigte COPELAND (1977) anhand von Simulationen, daß der Bias des relativen Risikos bei Fehldiagnosen in einer Kohortenstudie durch ein Absinken der Krankheitshäufigkeit verstärkt wird. Dieses Ergebnis wurde später von GREEN (1983) auch theoretisch hergeleitet. Analog dazu ist in einer Fall-Kontroll Studie der Bias bei Mißklassifikationen der Exposition um so größer, je extremer die Prävalenz wird.

### Mißklassifikationen und mehrdimensionale Kontingenztafeln

Die bei Vierfeldertafeln zu beobachtenden Effekte von Mißklassifikationen treten ebenso bei Kontingenztafeln höherer Ordnung auf. BROSS hatte für Vierfeldertafeln die Bedingungen dargestellt, die für einen gültigen asymptotischen Test auf Unabhängigkeit gelten müssen. Die Verallgemeinerung dieser Bedingungen für beliebig dimensionierte Kontingenztafeln geht auf KORN (1981) zurück. Korn untersuchte, welche hierarchischen log-linearen Modelle in der Gegenwart einer mißklassifizierten Kontingenztafel unverändert bleiben und somit die gleiche Struktur besitzen wie die unbeobachtbare wahre Kontingenztafel.

Für eine Einführung in log-lineare Modelle, siehe AGRESTI (1990). Zur Klärung der Nomenklatur sei das Beispiel eines saturierten log-linearen Modells für eine dreidimensionale Kontingenztafel angegeben:

$$\log \pi_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)}.$$

Notwendig, aber noch nicht hinreichend für die Bewahrung des log-linearen Modells unter Mißklassifikationen (in möglicherweise mehreren Dimensionen der Kontingenztafel), sind die folgenden zwei Bedingungen:

1. Für jede Dimension der Tafel muß die bedingte Wahrscheinlichkeit, daß bei einem gegebenem wahren Wert mißklassifiziert wird, unabhängig sein von den wahren Werten der Beobachtung in den anderen Dimensionen.
2. Bei gegebenen wahren Werten der Beobachtung muß die bedingte Wahrscheinlichkeit einer Mißklassifikation in einer bestimmten Dimension unabhängig von Mißklassifikationen in anderen Dimensionen sein.

Bevor das Theorem von Korn formuliert wird, ist die Definition eines *minimal set* hilfreich: Ein *minimal set* von  $u$ -Termen eines hierarchischen log-linearen Modells ist definiert als die kleinste Anzahl derjenigen  $u$ -Terme, die das Modell durch alle Terme niedrigerer Ordnung bestimmt. Insgesamt folgt der Satz:

Ein hierarchisches log-lineares Modell behält in Gegenwart von Mißklassifikationen in der Dimension  $l$  genau dann seine Struktur, wenn die Bedingungen 1. und 2. erfüllt sind und genau ein  $u$ -Term im minimal set ein  $l$  als Index hat. Das Ergebnis von Bross bezüglich des  $\chi^2$ -Tests ist in diesem Sinne ein Spezialfall des Theorems von Korn. Bross untersuchte das Modell der Unabhängigkeit zweier binärer Variablen, welches auch als ein log-lineares Modell äquivalent formuliert werden kann. Dieses log-lineare Modell ändert sich unter nichtdifferentiellen Mißklassifikationen nach dem Theorem von Korn nicht, so daß die auf dem Modell beruhenden Tests gültig bleiben.

### Mißklassifikationen und Trendtest

Auch beim Trendtest ist die Frage nach der Gültigkeit im Falle von Mißklassifikationen zu stellen. BRENNER (1992) und andere von ihm zitierte Autoren untersuchen dies für den Fall der nichtdifferentiellen Fehler bei einer ordinalen Expositionsvariablen. Der Autor stellt fest, daß ein Test auf Trend auch bei nichtdifferentiellen Fehlern zu einem erniedrigten  $p$ -Wert führen kann. Der *Attenuationseffekt* nichtdifferentieller Fehler (daß die Verzerrung zur Nullhypothese hin geschieht) muß hier nicht unbedingt auftreten.

Wenn die Beziehung zwischen Exposition und Zielvariable monoton ist, dann ist eine Antikonservativität des Tests eher unwahrscheinlich (aber nicht unmöglich). Ist sie es jedoch nicht, dann ist Antikonservativität nichts Ungewöhnliches.

Nichtdifferentielle Fehler zwischen zwei Expositionsstufen führen immer zu einem Attenuationseffekt, wenn beide Gruppen in diesen beiden Expositionsstufen miteinander verglichen werden. Bei einem Vergleich von einer dieser beiden Gruppen mit einer dritten Expositionsausprägung können jedoch nichtdifferentielle Fehler zu verzerrten Schätzungen in beiden Richtungen, sowie zur Antikonservativität des Tests führen. Eine Ausnahme in diesem Fall bildet der Vergleich der niedrigsten mit der höchsten Expositionskategorie. Dieser kann bei einem linearen Trend nur zur Nullhypothese hin verzerrt sein, so daß der Test zwar an Power verliert, aber niemals antikonservativ werden kann. Zusätzlich können nichtdifferentielle Fehler auch die Form des Zusammenhangs ändern (z. B. lineare-nichtlineare Beziehung).

Wenn ein linearer Trend vorliegt, dann kann bei nichtdifferentiellen Fehlern innerhalb benachbarter Expositionsstufen von einem Attenuationseffekt ausgegangen werden. Notwendig bleibt jedoch für diese Aussage, daß Mißklassifikationen nur zu benachbarten Kategorien hin stattfinden. Im allgemeinen ist es jedoch nicht bekannt, ob die wahren, nicht beobachteten Daten einen Trend aufweisen, so daß zur Abschätzung der Richtung der Verzerrung die a-priori Annahme eines linearen Trends nicht gerechtfertigt ist.

Tabelle 2: Effekte nichtdifferentieller Fehler beim Test auf Trend

	Wahre Kontingenztafel		Beobachtete Kontingenztafel	
	Kolorektalkarzinom ja	nein	Kolorektalkarzinom ja	nein
hoher OGK	35	40	28	32
normaler OGK	10	25	17	33
niedriger OGK	55	35	55	35
$\Sigma$	100	100	100	100
Odds-ratio	0,74		0,72	
p-Wert	0,051		0,047	

Tabelle 2 veranschaulicht, daß bereits für nichtdifferentielle Fehler der Test auf Trend ungültig werden kann und Schätzungen von der Nullhypothese weg verzerrt werden können. Dabei wird angenommen, daß in der Expositionsstufe „hoher Obst- und Gemüsekonsum (OGK)“ 20% der Studienteilnehmer ihren Konsum als „normal“ angeben. Als Resultat fällt der p-Wert und das Odds-ratio zu niedrig aus, und die Testentscheidung wird auf dem 5%-Niveau umgedreht.

Auch vor dem Hintergrund des Meßfehlerproblems zeigt sich, daß es wenig sinnvoll ist, Variablen simplifizierend zu dichotomisieren. HOWE & CHIARELLI (1988) stellen fest, daß der Powerverlust bei nichtdifferentiellen Fehlern durch die Bildung einer Vierfeldertafel maximiert wird. Der Gewinn an Trennschärfe ist markant, wenn (anstatt zwei) drei oder besser noch vier Kategorien gebildet werden (zusätzliche Kategorien bringen nur noch geringere Steigerungen der Power). Denn bei mehreren Kategorien können Fehlklassifizierungen partiell dadurch ausgeglichen werden, daß ein Teil „nur“ in benachbarte Kategorien fehlerklassifiziert werden. Dennoch kann es natürlich aus anderen Gründen, z. B. bei Fallzahlproblemen, sinnvoll sein zu dichotomisieren.

Wenn in einer Kontingenztafel mit mehr als zwei Expositionsstufen nichtdifferentiell fehlerklassifiziert wird, dann kann eine Kontingenztafel mit neudefinierter binärer Exposition differentielle Fehler aufweisen (BRENNER & BLETTNER, 1993, vgl. auch FLEGAL et al., 1986).

### Meßfehler im Confounder

Nicht nur Fehler in der Expositionsbestimmung verzerren die Schätzung des Expositionseffektes, sondern auch fehlerbehaftete Confounder. GREENLAND (1980) untersuchte den Fall nichtdifferentieller Fehler bei der  $2 \times 2 \times 2$ -Tafel. Dabei stellte er fest, daß Confoundermeißklassifikationen zu einer verzerrten Schätzung des Expositionseinflusses führen, da das Confounding nicht mehr vollständig kontrolliert werden kann. Somit liegt das beobachtete Odds-ratio der Exposition zwischen dem nach dem Confounder unadjustierten und dem vollständig adjustierten Wert. Dies gilt auch, wenn der Confounder mehr als nur zwei Ausprägungen und eine lineare Beziehung zum Erkrankungsrisiko besitzt. Wenn sich diese Assoziation jedoch als nicht-monoton darstellt, dann kann das adjustierte Odds-ratio noch weiter von dem wahren Wert entfernt liegen als das nicht-adjustierte Odds-ratio (BRENNER 1998).

SAVITZ & BARON (1989) geben den Bias des Expositions-Odds-ratios als eine Funktion von Sensitivität und Spezifität der Confoundermessungen an (nichtdifferentielle Fehler). Dabei zeigen sie, daß sogar mit einer relativ hohen Sensitivität und Spezifität von z. B.  $Se = Sp = 0,95$  ein starker Verzerrungseffekt verbunden sein kann.

Tabelle 3: Effekte nichtdifferentieller Fehler eines Confounders

	Wahre Kontingenztafel		Beobachtete Kontingenztafel	
	Fall	Kontrolle	Fall	Kontrolle
Confounder = 0				
Exponiert	20	10	100	18
Nicht exponiert	100	50	140	54
Confounder = 1				
Exponiert	400	40	320	32
Nicht exponiert	200	20	160	16
$\Sigma$	720	120	720	120
Odds-ratio(Exp)	1		1,5	
p-Wert	1		0,05	

MARSHALL & HASTRUP (1996) stellten fest, daß bei nichtdifferentiellen Fehlern eine Exposition, die keinen Einfluß auf die Krankheit hat, aufgrund einer negativen Korrelation mit dem Confounder protektiv erscheint, wenn der Confounder ein Risikofaktor ist und mißklassifiziert wird. Umgekehrt würde bei einer positiven Korrelation die Exposition fälschlicherweise als schädlich angesehen werden. Diese Effekte werden um so stärker, je stärker der Zusammenhang von Confounder und Exposition, je größer der Effekt des Confounders und je schlechter die Meßmethode ist.

Das Odds-ratio der Exposition wird stärker verzerrt, wenn die Sensitivität von 1 auf 0,95 fällt, als wenn sie beispielsweise von 0,6 auf 0,55 fallen würde. Marshall und Hastrup zeigen, daß eine Sensitivität und Spezifität von 0,95 ein Odds-ratio ergeben kann, das nur ungefähr die Hälfte des wahren Odds-ratios ist. Dieses Ergebnis ist für den Anwender beunruhigend, weil in der Praxis die Beobachtungsfehler häufig noch wesentlich größer sind.

In der Tabelle 3 wird angenommen, daß der Confounder C in 20% aller Fälle nichtdifferentiell mißklassifiziert wird, wenn  $C = 1$  gilt. Die Exposition hat keinen Einfluß auf die Krankheit, also  $OR(Exp) = 1$ . Durch die Mißklassifikation steigt die Heterogenität der beiden Expositions-Odds-ratios. Für  $C = 0$  erhöht sich das geschätzte Odds-ratio auf 2,14, während es für  $C = 1$  bei 1 bleibt. Insgesamt, adjustiert nach dem Confounder, ergibt sich ein geschätztes Expositions-Odds-ratio von 1,5, wobei der entsprechende p-Wert von 1 auf 0,01 fällt. Der p-Wert erniedrigt sich weiter auf 0,01, wenn eine Confounder-Expositions-Interaktion mit in das Modell aufgenommen wird.

Insgesamt ist aus diesen Analysen ersichtlich, daß eine Confundermißklassifikation noch größeren Schaden anrichten kann als eine fehlerhafte Bestimmung der primär interessierenden Einflußgröße (KUPPER, 1984). Natürlich wird aus der Mißklassifikation eines schwachen Confounders auch nur ein schwacher Bias resultieren, aber in der Epidemiologie muß oft nach sehr starken Risikofaktoren, wie etwa *Rauchen*, adjustiert werden.

Generelle Aussagen zum Verhalten der Odds-ratios bei differentiellen Mißklassifikationen lassen sich nicht treffen, da sie u. a. auf der gemeinsamen Verteilung der Variablen beruhen (GREENLAND, 1980).

Wie im Beispiel gesehen, beeinflußt die Mißklassifikation des Confounders zusätzlich auch die Heterogenität bzw. die Homogenität der Odds-ratios über die einzelnen Strata (GREENLAND, 1980). Es kann eine Heterogenität der Odds-ratios übersteigert als auch überdeckt werden. Inwieweit Confundermißklassifikationen für die Heterogenität der

beobachteten Odds-ratios verantwortlich sein können, wird z. B. von WALKER & LANES (1991) untersucht. Die Autoren bauen auf Arbeiten von BLETTNER & WAHRENDORF (1984), sowie von COX & ELWOOD (1991) auf, um bedingt auf den beobachteten Häufigkeiten zulässige Mißklassifikationsproportionen zu erhalten. Mit diesen kann dann untersucht werden, ob die Heterogenität der Odds-ratios durch Mißklassifikationen des Confounders erklärbar ist. Walker und Lanes geben ein Beispiel an, in dem die Heterogenität zweier Odds-ratios ( $OR_1 = 1,0$ ,  $OR_2 = 6,5$ ) untersucht wird. Dabei finden sie heraus, daß in diesem Fall eine nichtdifferentielle Fehlklassifizierung des Confounders die Heterogenität sogar noch verstärken würde und mithin nicht zur Erklärung der Heterogenität herangezogen werden kann. Bei differentieller Mißklassifizierung bezüglich Y sind verschiedene Szenarien denkbar, z. B. daß die Kontingenztafel das Ergebnis von  $OR_1 = OR_2 = 0,5$  bei entsprechenden Fehlklassifizierungen darstellt.

### Meßfehler bei der multiplen logistischen Regression

Die Effekte von Meßfehlern in den Variablen im logistischen Regressionsmodell sind von einer Vielzahl von Faktoren abhängig, vergleiche dazu ARMSTRONG (1990). Wird eine logistische Beziehung zwischen der wahren Variablen  $X$  und des Fall-Kontroll-Status  $D$  und ein nichtdifferentielles Meßfehlermodell (z. B. der Form  $Z = X + E$ ) vorausgesetzt, dann folgt auch der Zusammenhang zwischen der fehlerbehafteten Variablen  $Z$  und  $D$  der logistischen Funktion. Bei differentiellen Fehlern würde dieses Ergebnis jedoch nicht bestehen bleiben. Ebenso wie der  $\chi^2$ -Test bei einer Kontingenztafel bleibt im univariaten Fall der Test auf Assoziation bei nichtdifferentiellen Fehlern gültig, wobei wiederum ein Powerverlust zu konstatieren ist. Ebenfalls analog zur Vierfeldertafel ist die Effektschätzung zur Nullhypothese hin verzerrt. Im Falle des Meßfehlermodells  $Z = X + E$  mit unabhängig identisch normalverteilten Meßfehlern mit Erwartungswert Null kann auch das Maß der Effektverzerrung angegeben werden. Für den auf  $Z$  beru-

henden verzerrten Schätzer  $\beta_Z$  gilt dann  $\beta_Z = \frac{\sigma_X^2}{\sigma_Z^2} \cdot \beta_X$ . Damit wird deutlich, daß die

Effektschätzung um so stärker verzerrt, je größer der Meßfehler ist. Da  $Z$  variabler als  $X$  ist, wird  $\beta_Z$  neben der Verzerrung auch weniger präzise geschätzt. Ein entsprechendes Konfidenzintervall würde also breiter sein als das von  $\beta_X$ . Andererseits ist es eventuell nicht breit genug, weil es aufgrund des Bias um einen Wert zentriert ist, der evtl. eine große Entfernung zum wahren Expositionsrisiko aufweist.

Der Attenuationseffekt ist von dem vorliegenden Meßfehlermodell abhängig. Gilt nicht das obige klassische Meßfehlermodell sondern das *Berksonmodell*  $E(X|Z) = Z$ , dann ist keine Verzerrung zur Nullhypothese hin zu beobachten (ARMSTRONG, 1990).

Wenn im multivariaten Fall die unverfälschten Variablen einer multivariaten Normalverteilung folgen und die Meßfehler in den Expositionsbestimmungen unabhängig von allen anderen Kovariablen und ebenfalls normalverteilt sind, dann ist eine zum univariaten Fall analoge Berechnung des Bias möglich. Wenn zusätzlich Confounder in dem logistischen Regressionsmodell berücksichtigt werden, dann ist bei einem identischen Meßfehler in der interessierenden Exposition eine stärkere Verzerrung zu beobachten. Diese bleibt sogar bestehen, wenn der Confounder keinen Effekt auf die Krankheit hat und somit unnötigerweise in das Modell mit aufgenommen wurde.

Wie bei entsprechenden Kontingenztafeln im kategoriellen Fall gilt auch im stetigen Fall, daß ein meßfehlerbehafteter Confounder bei nichtdifferentiellen Fehlern das Confounding höchstens partiell kontrolliert und daß die partiell kontrollierte Effektschätzung bei einer monotonen Beziehung des Confounders zum Erkrankungsrisiko zwischen dem vollständig kontrollierten und dem nicht kontrollierten Wert liegt. In diesem Fall kann nicht davon ausgegangen werden, daß der Test auf Assoziation gültig ist. Sobald also



Tabelle 4: Zusammenfassung wichtiger Meßfehlereffekte

	Schätzungen Richtung der Verzerrung		Tests Gültig: ja/nein	
	Nichtdiff. Fehler	Diff. Fehler	Nichtdiff. Fehler	Diff. Fehler
Kategorielle Exposition				
Vierfeldertafel	$H_0$	Offen	Ja, Powerverlust	Nein
Trendtest	Offen	Offen	Nein	Nein
Meßfehler im Confounder	Offen	Offen	Nein	Nein
Stetige Exposition				
Klass.Meßf.modell: $E(Z) = X$	$H_0$	Offen	Ja, Powerverlust	Nein
Berksonmodell: $E(X) = Z$	Unverzerrt	Offen	Ja, Powerverlust	Nein
Meßfehler im Confounder	Offen	Offen	Nein	Nein

eine einzige Variable im multivariaten logistischen Regressionsmodell Meßfehlern unterworfen ist, können die mit dieser Variablen korrelierten Expositionen nicht mehr unverzerrt bewertet und allgemein gültigen Tests unterzogen werden.

Existiert zusätzlich zum fehlerbehafteten Confounder ein Fehler in der Exposition, dann verbinden sich beide Verzerrungen, wobei bei korrelierten Fehlern in Exposition und Confounder die Richtung und der Betrag der Verzerrung verändert werden können.

Der Bias der Expositionsschätzungen wird um so größer, je höher der Meßfehler und je größer der Betrag der Schätzung ist. Im kategoriellen Fall ist der Bias mit der Prävalenz der Exposition assoziiert. Die aufgezeigten Resultate gelten auch für nichtdifferentiell mißklassifizierte kategorielle Variablen.

Tabelle 4 faßt die wichtigsten Resultate der Meßfehlereffekte bezüglich Schätzungen und Tests schlagwortartig zusammen. „Ungültig“ bei Tests bedeutet dabei, daß *im allgemeinen* die Gültigkeit des Tests nicht angenommen werden darf. Es ist jedoch möglich, daß je nach Art der Mißklassifikation auch ihr Niveau haltende Tests resultieren. Ein gültiger Test dagegen hält immer sein Niveau, wenngleich dieses nicht voll ausgeschöpft werden muß.

## Diskussion

Die Effekte fehlerhafter Messungen sind nicht auf den Bereich der Epidemiologie beschränkt. Die Verwendung von ungenauen Zielgrößen oder Surrogatparametern führt z. B. auch bei kontrollierten klinischen Therapiestudien oder bei Anwendungsbeobachtungen (RIPPIN & MICHAELIS, 1998) zu gleichen Problemen, wenngleich in diesem Bereich der medizinischen Forschung die Auswirkungen unscharfer Messungen bisher weniger beachtet wurden. Im Bereich randomisierter und doppelblinder Studien ist bei einem Meßfehler in der Zielgröße i. a. von einem nichtdifferentiellen Fehler auszugehen. Therapieunterschiede wären dann zur Nullhypothese hin verzerrt und die Power der Tests wäre erniedrigt. Auch ein Fehler in einer z. B. dichotomen Baseline-Variablen würde eine Gegenüberstellung der Therapieresponse für diese Variable verzerren und die Power des Vergleichs erniedrigen.

In diesem Artikel wurden die einzelnen Methoden zur Korrektur dieser Art von Verzerrungen und auch die Designmöglichkeiten und Fallzahladjustierungen nicht diskutiert. Durch die Konzeption der Validierungsstudien oder der Replikationsstudien können Meßfehlermodelle aufgestellt werden und gültige Tests und unverzerrte Schätzungen ent-

wickelt werden. Eine interne Validierungsstudie ist eine Substichprobe der Hauptstudie, in der das interessierende Merkmal zusätzlich fehlerfrei bestimmt wird, während bei einer Replikationsstudie durch wiederholte Messungen die Ungenauigkeit in den Messungen abgeschätzt wird. Diese Konzepte sollten in der Regel im Vorhinein einer Studie in die Designüberlegungen eingehen, obwohl auch nach der Durchführung einer Studie die Validierung eines Merkmals sinnvoll sein kann. Eine sequentielle Studienplanung, in der zuerst nur eine begrenzte, komplett validierte Stichprobe erhoben wird, um die Notwendigkeit oder kostenoptimale Größe (BUONACCORSI (1990), GREENLAND (1988), PALMGREN (1987)) der endgültigen Validierungsstudie festzulegen, kann zu einem (kosten-) optimierten Design führen.

Eine bloße Erhöhung der Fallzahl, um den Powerverlust bei nichtdifferentiellen Fehlern auszugleichen, greift in der Regel zu kurz, da eine Verzerrung der Schätzung bestehen bleiben würde. Die Epidemiologie ist aber gerade daran interessiert, Risiko- oder protektive Faktoren *quantifizieren* zu können. Es reicht nicht aus, lediglich festzustellen, daß eine Exposition einen schädlichen oder auch günstigen Effekt auf eine Krankheit ausübt, ohne die Größenordnung dieses Einflusses abschätzen zu können.

Bereits bei der Studienplanung ist es wichtig, die Effekte fehlerhafter Variablen zu bedenken, um potentielle Designmaßnahmen zu treffen. Aber auch bei der im Nachhinein stattfindenden Beurteilung einer Studie muß überprüft werden, ob durch ungenaue Messungen Verzerrungen stattgefunden haben und wenn ja, in welche Richtung diese zu vermuten sind.

## Literatur

- AGRESTI, A. (1990): Categorical data analysis, Wiley, New York.
- ARMSTRONG, B. G. (1990): The effects of measurement errors on relative risk regressions, *American Journal of Epidemiology* **132**, 1176–1184.
- ASSAKUL, K., PROCTOR, C. H. (1967): Testing independence in two-way contingency tables with data subject to misclassification, *Psychometrika* **32**, 67–76.
- BLETTNER, M., WAHRENDORF, J. (1984): What does an observed relative risk convey about possible misclassification?, *Methods of Information in Medicine* **23**, 37–40.
- BRENNER, H. (1992): Notes of the assessment of trend in the presence of nondifferential exposure misclassification, *Epidemiology* **3**, 420–427.
- BRENNER, H. (1998): A potential pitfall in control of covariates in epidemiologic studies, *Epidemiology* **9**, 68–71.
- BRENNER, H., BLETTNER, M. (1993): Misclassification bias arising from random error in exposure measurement strategies, *American Journal of Epidemiology* **138**, 453–461.
- BROSS, I. D. J. (1954): Misclassification in 2×2 tables, *Biometrics* **10**, 478–486.
- BUONACCORSI, J. P. (1990): Double Sampling for exact values in the normal discriminant model with application to binary regression, *Communications in Statistics A* **19**, 4569–4586.
- COPELAND, K. T., CHECKOWAY, H., MCMICHAEL, A. J., HOLBROOK, R. H. (1977): Bias due to misclassification in the estimation of relative risk, *American Journal of Epidemiology* **105**, 488–495.
- COX, B., ELWOOD, J. M. (1991): The effect on the stratum-specific odds ratio of non-differential misclassification of a dichotomous covariate, *American Journal of Epidemiology* **133**, 202–207.
- FLEGAL, K. M., BROWNIE, C., HAAS, J. D. (1986): The effects of exposure misclassification on estimates of relative risk, *American Journal of Epidemiology* **123**, 736–751.
- GREEN, M. (1983): Use of predictive values to adjust relative risk estimates biased by misclassification of outcome status, *American Journal of Epidemiology* **117**, 98–105.
- GREENLAND, S. (1980): The effect of misclassification in the presence of covariates, *American Journal of Epidemiology* **112**, 564–569.
- GREENLAND, S. (1988): Statistical uncertainty due to misclassification: implications for validation substudies, *Journal of Clinical Epidemiology* **41**, 1167–1174.
- GULLEN, W. H., BEARMAN, J. E., JOHNSON, E. A. (1968): Effects of misclassification on epidemiology

- logic studies, Public Health Reports **83**, 914–918.
- HOLFORD, R. H., STACK, C. (1995): Study design for epidemiologic studies with measurement error, Statistical Methods in Medical Research **4**, 339–358.
- HOWE, R. H., CHIARELLI, A. M. (1988): Methodological issues in cohort studies 2: Power calculations, International Journal of Epidemiology **2**, 464–468.
- KEYS, A., KIHLEBERG, J. E. (1963): Effect of misclassification on estimated relative prevalence of a characteristic: Part 1. Two populations infallibly distinguished; Part 2. Errors in two variables, American Journal of Public Health **53**: 1656–1665.
- KORN, E. L. (1981): Hierarchical log-linear models not preserved by classification error, Journal of the American Statistical Association **76**, 110–113.
- KUPPER, L. L. (1984): Effects of the use of unreliable surrogate variables on the validity of epidemiologic research studies, American Journal of Epidemiology **120**, 643–648.
- MARSHALL, J. R., HASTRUP, J. L. (1996): Mismeasurement and the resonance of strong confounders: uncorrelated errors, American Journal of Epidemiology **143**, 1069–1078.
- MOTE, V. L., ANDERSON, R. L. (1965): An investigation of the effect of misclassification on the properties of  $\chi^2$ -tests in the analysis of categorical data, Biometrika **52**, 95–109.
- NEWELL, D. J. (1963): Misclassification in 2x2 tables, Biometrics **19**, 187–188.
- PALMGREN, J. (1987): Precision of double sampling estimators for comparing two probabilities, Biometrika **74**, 687–694.
- RIPPIN, G., MICHAELIS, J. (1998): Referenz-kontrollierte Beobachtungsstudien – ein neues Konzept für Anwendungsbeobachtungen, Arzneimitteltherapie, **9**, 279–283.
- SAVITZ, D. A., BARON, A. E. (1989): Estimating and correcting for confounder misclassification, American Journal of Epidemiology **129**, 1062–1071.
- WALKER, A. M., LANES, S. F. (1991): Misclassification of covariates, Statistics in Medicine **10**, 1181–1196.

Eingang des Manuskripts: 18. 2. 99

Anschrift des Autors: Dipl.-Stat. Gerd Rippin, Institut für Medizinische Statistik und Dokumentation der Johannes Gutenberg-Universität Mainz, Langenbeckstr.1, 55131 Mainz. Tel.: (061 31) 17 67 08, email:rippin@imsd.uni-mainz.de

Diese Publikation enthält Teile der geplanten Dissertation des Autors.

## BUCHBESPRECHUNGEN/BOOKREVIEWS

YEARGERS, E. K.; SHONKWILER, R. W.; HEROD, J. V.  
**An Introduction to The Mathematics of Biology  
 With Computer Algebra Models**  
 1996, 434 S., DM 118,  
 ISBN 3-7643-3809-1  
 Birkhäuser Verlag AG, Basel

Die Autoren dieses Buches vertreten den Standpunkt, daß *mathematische Biologie* nicht die Übertragung von einer Wissenschaft in die andere ist, sondern eine eigene Einheit bildet. Gleichgültig, ob man dieser Überlegung nun bis zum letzten Punkt zustimmen kann, ist die Fülle der behandelten Beispiele für Biologen und Mathematiker in gleicher Weise interessant. Die Computersoftware *Maple* wird konsequent angewandt. Dadurch ist auch der Zugang für Leser mit geringeren Mathematikkenntnissen relativ leicht. Die umfangreiche, didaktisch gute Darstellung ist geeignet, das Verständnis für biologische Prozesse und Phänomene zu vertiefen.

H. Geidel, Stuttgart

BÖKER, F.

**S-PLUS Learning by Doing**

Eine Anleitung zum Arbeiten mit S-Plus

1997, 126 S. mit Diskette, DM 54,,

ISBN 3-8282-0049-4

Lucius & Lucius Verlagsges. Stuttgart

Das vorliegende Buch zum Arbeiten mit S-PLUS ist aus einem Skript für Praktikumsveranstaltungen entstanden. Dabei war es das Ziel, den Umgang mit der Statistik-Programmiersprache S-PLUS zu erlernen, eigene Funktionen zu schreiben und die Benutzung des Hilfesystems und des Handbuches zu üben. Dabei wurde viel Wert auf die Anwendung der graphischen Möglichkeiten von S-PLUS gelegt. Es sollte aber nicht nur das technische Lernen der Programmiersprache im Vordergrund stehen, sondern daneben sollte auch immer etwas Statistik gelernt werden, möglichst auch neuere Methoden, die zum Teil im Vorlesungsbetrieb noch zu kurz kommen.

Nutzen zieht man aus diesem didaktisch sehr gut aufgebauten und gestalteten Buch, wenn man sich an den Computer setzt und anfängt zu arbeiten. Die mitgelieferte Diskette ist dabei sehr hilfreich.

H. Geidel, Stuttgart

RIEDWYL, H.

**Lineare Regression und Verwandtes**

Beispiele mit Lösungsvorschlägen

1997, 144 S., DM 38,

ISBN 3-7643-5495-X

Birkhäuser Verlag AG, Basel

Eine Sammlung von Beispielen soll und kann das Verständnis für die Bearbeitung von statistischen Fragestellungen mit Hilfe der Regressionsanalyse schulen und vertiefen. Als Voraussetzung genügt eine einsemestrige Einführungsvorlesung in angewandter Statistik.

Das Buch ist in zwei Teile gegliedert: Im ersten Teil werden die Daten (im Internet abrufbar) unter Angabe der Quellen und einer kurzen Formulierung des Problems präsentiert. Im zweiten Teil werden Lösungsvorschläge diskutiert.

Neben der linearen Regressionsgeraden wird in Abhängigkeit von den Problemstellungen eine Vielzahl von weiteren Methoden angesprochen. Auf die wichtige Möglichkeit von Transformationen wird mehrfach hingewiesen. Für den Nachvollzug der Lösungen können gegebenenfalls vorhandene Programmsysteme eingesetzt werden.

Das vorliegende Buch kann durch die 41 Beispiele aus den verschiedensten Fachrichtungen das Wissen um die Anwendung der Regressionsanalyse vertiefen, aber auch viele Anregungen für die Lösung von analogen Problemen vermitteln.

H. Geidel, Stuttgart

CONWAY, J. H.; GUY, R. K.

**Zahlenzauber** Von natürlichen, imaginären und anderen Zahlen

Aus dem Amerikanischen von M. Stern

1997, 346 S., DM 58,, ISBN 3-7643-5244-2

Birkhäuser Verlag AG, Basel

Überall werden wir mit Zahlen konfrontiert. Jedoch machen wir uns selten Gedanken über die Art der Zahlen. Das vorliegende Buch will hier eine Hilfe sein. Dabei mag die Fülle der behandelten Zahlenfamilien, Zahlengruppen und Zahlenbereiche verwirrend aber auf der anderen Seite auch anregend wirken. Überraschend stellt man fest, daß hinter auf den ersten Blick unregelmäßig erscheinenden Zahlenfolgen doch letztlich Gesetzmäßigkeiten zu Tage treten. Ein klassisches Beispiel sind die Fibonaccischen Zahlen, auf die wir in der Natur immer wieder treffen. Hier ist ein Ansatz, sich intensiver mit Zahlensystemen zu beschäftigen. Dabei mag das gut ausgestattete Buch Hilfe und Anregungen vermitteln.

H. Geidel, Stuttgart

**Mitteilungen der  
Deutschen Gesellschaft für Medizinische Informatik,  
Biometrie und Epidemiologie (GMDS) e.V.**

**Inhalt nach Rubriken**

Brief des Präsidenten .....	Seite 1
Aufruf zur Gründung einer Projektgruppe „Methodik systematischer Reviews“ .....	Seite 2
BMBF-Zuschuß für die Antragstellung von EURO-Fördermitteln .....	Seite 3
Aus-, Fort- und Weiterbildung .....	Seite 3
Ankündigungen und Veranstaltungshinweise .....	Seite 6
Hinweis zur Mitgliedsbeitragsrechnung 1999 .....	Seite 8
Neuaufnahmen .....	Seite 9
Förderpreis für Studierende 1999 .....	Seite 9
Posterpreis 1999 .....	Seite 10
Briefwahl 1999 zum Präsidium der GMDS – Aufruf zur Wahl .....	Seite 10
Briefwahl 1999 zum Präsidium der GMDS – Kandidatenvorstellung .....	Seite 11
Ankündigung der ISCB-GMDS-Jahrestagung 1999 in Heidelberg .....	Seite 14
Vergünstigter Bezug der Zeitschrift „Technology and Health Care“ .....	Seite 16

**Brief des Präsidenten  
Prof. Dr. K.-H. Jöckel**

Liebe Kolleginnen und Kollegen,

gleich zu Anfang meines Schreibens möchte ich Sie auf folgendes besonders hinweisen: den Mitteilungen liegen die Wahlunterlagen zum Präsidium, für die Wahl zum 1. Vizepräsidenten, zum Schatzmeister und Schriftführer bei. Ich möchte Sie herzlich bitten, sich möglichst zahlreich an dieser Wahl zu beteiligen! Wie Sie wissen, liegt nach unserer Satzung das Vorschlagsrecht für die Wahl des Schatzmeisters beim Präsidium. Ich freue mich, Ihnen mitteilen zu können, daß es uns gelungen ist, Herrn Professor Dr. Schweim, Direktor des DIMDI, für diese Kandidatur gewinnen zu können. Im Namen des Präsidiums darf ich Sie bitten, ihm Ihre Stimme zu geben und ihm damit Unterstützung für die vor ihm liegenden Aufgaben zu signalisieren.

Zu den Tagungen ist folgendes zu sagen: die Vorbereitungen für Heidelberg sind im vollen Gang. Die Jahrestagung 2000 wird wie vorgesehen gemeinsam mit der EFMI in Hannover stattfinden.

In meinem letzten Brief hatte ich auf die durchgeführte Satzungsänderung hingewiesen. Leider zwingen Monita des Vereinsregistergerichts an einigen juristischen Formulierungen, uns auf der nächsten Mitgliederversammlung noch einmal kurz mit dieser Sache zu beschäftigen. Sie werden rechtzeitig die juristisch korrekten Formulierungen zugeleitet bekommen.

Auch wenn durch vielfache Aktivitäten des Bundes und der EU selber sehr deutlich auf die Fördermöglichkeiten im Rahmen des 5. Rahmenprogramms (1998-2002) aufmerksam gemacht wurden, möchte ich doch noch einmal darauf hinweisen, daß im Laufe dieses Jahres eine Reihe von Bewerbungsterminen für unsere Fächer relevant werden. Ihnen allen dürfte klar sein, wie wichtig es ist, daß sich möglichst viele Mitglieder der GMDS an diesen Ausschreibungen beteiligen. Einzelheiten zum 5. Rahmenprogramm können Sie unter <http://www.cordis.lu/fp5>, sowie national unter <http://www.kowi.de> einsehen.

Nun noch ein Hinweis für alle, die schon immer Ihren Antrag zur Erlangung eines der Zertifikate, die von unserer Fachgesellschaft mitvergeben werden, stellen wollten: Tun Sie dies noch in diesem Jahr, denn im nächsten Jahr steigen die Gebühren! Nachdem alle anderen beteiligten Fachgesellschaften zugestimmt haben, werden die Gebühren für die Erlangung des Zertifikats für Medizinische Informatik, Biometrie in der Medizin und Epidemiologie auf DM 200,-- angehoben, damit Sie in Zukunft kostendeckend sind.

Zum Abschluß komme ich der angenehmen Pflicht nach, den folgenden Kollegen, deren Geburtstag in das 2. Quartal dieses Jahres fällt, im Namen der GMDS recht herzlich zu gratulieren.

60. Geburtstag: Dr. Hans-Georg Eisenwiener, Basel



65. Geburtstag: Prof. Dr. Wilhelm Gaus, Ulm  
 Dr. rer. nat. Dipl. Phys. Karl-Josef Johannes, Köln  
 Dipl.-Ing. Till-R. Kornemann, Berlin  
 Herr Friedrich Linneweber, Köln  
 Dr. med. Uwe Beneicke, Duisburg  
 Dr. Gottfried Enderlein, Berlin

Mit freundlichen, kollegialen Grüßen  
 Ihr

Prof. Dr. K.-H. Jöckel  
 Präsident

**Aufruf zur Gründung einer neuen Projektgruppe  
 „Methodik systematischer Reviews“**

Im Bereich der Medizin nimmt die Bedeutung systematischer Übersichtsarbeiten immer mehr zu. In vielen Ländern wurde in den letzten Jahren verstärkt auch an vielfältigen methodisch ungeklärten Aspekten gearbeitet. Um einerseits in Deutschland methodische Forschungsarbeiten zu unterstützen und die Kontakte zwischen Wissenschaftlern zu vertiefen, andererseits aber auch um die medizinischen Kollegen bei der Durchführung systematischer Reviews unterstützen zu können, beabsichtigt der *Fachbereich Biometrie der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS)* die Gründung einer Projektgruppe

**“Methodik systematischer Reviews”**

Wir möchten hiermit alle interessierten Kolleginnen und Kollegen zur Gründungssitzung am

**14. Juni 1999 von 11.00 - 17.00 Uhr**

Hotel Kolpinghaus (ca. 800 m vom Hbf), Goethestraße 13, **Fulda**,

einladen.

Es wird zwei Sitzungen mit Vorträgen zu methodischen Aspekten systematischer Reviews und eine Diskussionsrunde zu den Aufgaben und Zielen der zukünftigen Projektgruppe geben. Es sollen im Anschluß Wahlen der Leiterinnen und Leiter der Projektgruppe durchgeführt werden. Die nächste Sitzung dieser Projektgruppe könnte anlässlich der GMDS-Tagung in Heidelberg stattfinden (13. - 16. Sept. 1999).

Je ein Einführungs-/ Übersichts Vortrag zu systematischen Reviews bei randomisierten Studien und bei Beobachtungsstudien werden von Dr. G. Antes (Freiburg) und Prof. M. Blettner (Bielefeld) gehalten.

Weitere Vortragsanmeldungen sind bis **31.05.1999** möglich. Mitteilung über die Annahme erfolgt am 02. Juni per Email. Formlose Anmeldung mit Email-Adresse bis **07. Juni** erwünscht.

Anmeldung von Vorträgen und weitere Auskünfte bei:

Prof. Dr. Markus Löffler, Institut für Med. Informatik, Statistik und Epidemiologie, Univ. Leipzig,  
 Liebigstr. 27, 04103 Leipzig, Tel.: 0341/9716100, Fax: 0341/9716109,  
 Email: loeffler@imise.uni-leipzig.de

und

Informatik, Biometrie und Epidemiologie in Medizin und Biometrie - Band 30 - Heft 2/1999



Dr. Willi Sauerbrei, Institut für Med. Biometrie und Med. Informatik, Uni-Klinikum Freiburg, Stefan-Meier-Straße 26, 79104 Freiburg, Tel.: 0761/203-6669, Fax: 0761/203-6677, Email: wfs@imbi.uni-freiburg.de

P.S. Unter dem Stichwort 'GMDS' können im Kolpinghaus Zimmer zum Sonderpreis (incl. Frühstück EZ DM 90,-, DZ 140,- unter Tel.: 0661/8650-0) gebucht werden.

### **BMBF-Zuschuß für die Antragstellung von EURO-Fördermitteln**

Das BMBF gewährt bis zu DM 10.000 Zuschuß für die Antragstellung von EURO-Fördermitteln. Damit soll der Mehraufwand bei der Antragstellung für Forschungsförderungsmittel der EU aus dem Bereich Lebenswissenschaften (BIOMED2-Nachfolgeprogramm) für Anträge, die in 1999 bei der EU eingereicht werden sollen, aufgefangen werden. Das BMBF gewährt o.g. Zuschuß für organisatorischen Mehraufwand bei Porto-, Telefon- und FAX-Kosten, Reisemittel, Büromaterial, Aufträge an Dritte und zusätzlichen Personalaufwand. Ein Rechtsanspruch auf die Unterstützung besteht nicht. Nähere Informationen und ein Musterblatt für die Antragstellung erhalten Sie unter: DLR Projektträger des BMBF, Programmkoordination Lebenswissenschaften, Königswinterer Str. 522-524, 53227 Bonn, FAX: (0228) 4492-333. Alle Unterlagen für die Antragstellung bei der EU können über das Internet bezogen werden: <http://www.cordis.lu/life/calls/199901.htm>.

### **Aus-, Fort- und Weiterbildung**

#### **Veranstaltungen der Akademie Medizinische Informatik**

1. Seminar "Verschlüsselung von medizinischen Diagnosen und Therapien" - praktische Übungen  
Termin, Ort: 18.6.99, DKFZ, Heidelberg  
Zielgruppe: Ärzte, Medizinische Dokumentare und anderes mit Verschlüsselung betrautes Personal in Krankenhäusern, Arztpraxen und wissenschaftlichen Instituten sowie bei den Krankenkassen.  
Referenten: Dr. med. Bernd Graubner, ZI Köln, Prof. Dr. med. Rudolf Thurmayer, Universität München.  
Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS 320.- DM  
(Sonstige 800.- DM, Studierende 160.- DM)  
Hinweis: *Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*
  
2. Seminar "Patienteninformation in der Arztpraxis und im Krankenhaus"  
Termin, Ort: 19.6.99, DKFZ, Heidelberg  
Zielgruppe: Das Seminar wendet sich an Krankenhäuser, die sich zum Thema „Serviceleistung Patienteninformation“ informieren wollen..  
Referenten: Dipl.-Inform. Med. Michael Hägele, Dipl.-Inform. Med. Nela Sljivljak, Heidelberg  
Teilnahmegebühren: Mitglieder von GI, BVMI, DVMD, GMDS 320.- DM  
(Sonstige 800.- DM, Studierende 160.- DM)
  
3. Tutorium "Medizin-Controlling und dessen Unterstützung durch Informationstechnologie"  
Termin, Ort: 12.9.99, im Rahmen der internationalen Tagung ISCB-GMDS-99, Heidelberg  
Zielgruppe: Ärzte, Medizininformatiker, Pflegekräfte und Medizinische Dokumentare, die im Bereich des Medizin-Controlling tätig sind oder damit zu tun haben bzw. beruflich planen. Interessant ist das Tutorium auch für Mitarbeiter der Medizinischen Informatik im all-



gemeinen, im ärztlichen und Pflegedienstbereich, im Patientenmanagement sowie im Controlling, der Kosten-Leistungsrechnung und innerhalb der Krankenhaus-Betriebsleitung.

Referent: Univ.-Prof. Dr. Andreas J. W. Goldschmidt, Universität Bonn

Teilnahmegebühren: Mitglieder von GI, BVMI, DVMD, GMDS **300.- DM**

(Sonstige 450.- DM, Studierende 200.- DM)

*Hinweis: Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

#### **4. Seminar "Kodierungssoftware für Diagnosen und Prozeduren im Vergleich"**

Termin, Ort: 17.9.99, Kopfklinik, Heidelberg

Zielgruppe: Anbieter und Anwender von KIS-Systemen

Referenten: Dr. Josef Ingenerf, Universität Lübeck; PD Dr. med. Roswitha Thurmayr, TU München

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS **320.- DM**

(Sonstige 800.- DM, Studierende 160.- DM)

*Hinweis: Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

#### **5. Seminar "Analyse, Auswahl und Einführung von Krankenhausinformationssystemkomponenten"**

Termin, Ort: 7.10.99, Leipzig

Zielgruppe: Das Seminar wendet sich an Personen in Krankenhäusern, die mit der Einführung von Komponenten für Krankenhausinformationssysteme befaßt sind. Hierzu gehören insbesondere Abteilungsleiter, Projektleiter oder alle Personen, die Kenntnisse über das taktische Management von Krankenhausinformationssystemen erwerben oder ihr bisheriges Wissen aktualisieren wollen. Darüber hinaus richtet sich das Seminar auch an interessierte Personen bei Beratungsunternehmen bzw. Anbietern von Komponenten für Krankenhausinformationssysteme.

Referenten: Prof. Dr. sc. hum. Alfred Winter, Dipl.-Math. Gabriele Herrmann, Dr. oec. habil. Alfred Scharsky, Dr. sc. hum. Birgit Schneider, Universität Leipzig

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS **320.- DM**

(Sonstige 800.- DM, Studierende 160.- DM)

*Hinweis: Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

#### **6. Seminar "Einführung von Komponenten für Klinische Arbeitsplatzsysteme - Probleme, Hemmnisse und Lösungsstrategien"**

Termin, Ort: 8.10.99, Universität Leipzig

Zielgruppe: Das Seminar wendet sich an Personen in Krankenhäusern, die mit der Planung, Einführung und Betreuung von Komponenten für Klinische Arbeitsplatzsysteme beauftragt sind. Hierzu gehören insbesondere Medizininformatiker, Informatiker, Ärzte, Projektleiter im Rahmen von Einführungsprozessen, Mitarbeiter des Pflegedienstes etc. Darüber hinaus richtet sich das Seminar auch an interessierte Personen bei Beratungsunternehmen bzw. Anbietern von Komponenten für Krankenhausinformationssysteme.

Referentin: Dipl.-Math. Gabriele Herrmann, Universität Leipzig

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS **320.- DM**

(Sonstige 800.- DM, Studierende 160.- DM)

#### **7. Seminar "Gesundheitsökonomie"**

Termin, Ort: 11.-12.10.99, DKFZ, Heidelberg

Zielgruppe: Personen, die in die Planung und Durchführung gesundheitsökonomischer Studien involviert bzw. an Entscheidungsprozessen beteiligt sind, und die das Ziel einer (verbesserten) wirtschaftlichen Gesundheitsversorgung von Patienten verfolgen.



Personen, die an der Erfassung, Verwaltung und Verarbeitung von Kosten- und Leistungsdaten für Controlling und zum Qualitätsmanagement arbeiten.

Ärzte, die bei der Therapiewahl auch ökonomische Informationen bewerten bzw. einbeziehen möchten.

Personen, die mehr über Kosten- und Nutzenbewertung im Gesundheitswesen wissen möchten.

**Spezifische Vorkenntnisse werden nicht vorausgesetzt.**

Referent: Dipl.-Inform. Med. Oliver Mast, Leverkusen

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS 500.- DM

(Sonstige 1.200.- DM, Studierende 300.- DM)

Hinweis: *Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

#### **8. Seminar "Verschlüsselung von Diagnosen und Operationen mit ICD-9/10 und OPS"**

Termin, Ort: 13.10.99, DKFZ, Heidelberg

Zielgruppe: Ärzte, Medizinische Dokumentare, Medizinische Informatiker

Referent: Dr. med. Albrecht Zaiss, Universität Freiburg

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS 320.- DM

(Sonstige 800.- DM, Studierende 160.- DM)

Hinweis: *Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

#### **9. Seminar "Medizin für Nicht-Mediziner"**

Termin, Ort: 14.-15.10.99, RWTH Aachen

Zielgruppe: Informatiker, Medizininformatiker, Studierende der Medizinischen Informatik und der Informatik

Referent: Prof. Dr. Dr. Klaus Spitzer, RWTH Aachen

Teilnahmegebühren: Mitglieder von BVMI, DVMD, GMDS 500.- DM

(Sonstige 1.200.- DM, Studierende 300.- DM)

Hinweis: *Dieses Seminar kann für den Erwerb des gmds-Zertifikats 'Medizinische Informatik' angerechnet werden.*

\*\*\*

### **Postgraduelle Ausbildung MEDIZINISCHE BIOMETRIE**

Die Medizinische Fakultät der Universität Heidelberg bietet, unter Federführung des Instituts für Medizinische Biometrie und Informatik, ein postgraduelles Ausbildungsprogramm in Medizinischer Biometrie an. Die Ausbildung kann flexibel in thematisch abgeschlossenen Blöcken mit individueller Kurswahl absolviert werden, so daß eine berufsbegleitende Weiterbildung möglich ist. Bei erfolgreichem Durchlaufen eines definierten Curriculums führt das Programm zu einem Universitätszertifikat mit Fachanerkennung der GMDS. Die Kurse können auch einzeln besucht werden, ohne daß die Erlangung des Zertifikats angestrebt wird.

In der zweiten Jahreshälfte 1999 werden in Heidelberg die folgenden Kurse angeboten. Dabei werden die genannten Kursleiter durch weitere Dozenten unterstützt.

#### **GRUNDLAGENKURS MEDIZIN**

15.07. - 20.07.1999 Einführung  
Prof. Dr. R. Nobiling, Universität Heidelberg

08.10. - 10.10.1999 Bewegungsapparat / Orthopädie  
Dr. K.-L. Krämer, Universität Heidelberg



**GRUNDLAGENKURS STATISTIK**

Prof. Dr. F.-T. Nürnberg, FH Mannheim

19.08. - 23.08.1999, 02.09. - 06.09.1999, 09.09. - 13.09.1999

**AUFBAUKURSE**

11. - 13.11.1999      Überlebenszeitanalyse  
 Prof. Dr. M. Schumacher, Universität Freiburg

**Wahlkurse**

17. - 19.06.1999      Methoden der klinischen Pharmakologie  
 Prof. Dr. M. Wehling, Universität Heidelberg  
 Prof. Dr. W. E. Haefeli, Universität Heidelberg

24.06. - 26.06.1999      Evaluation medizinischer Diagnoseverfahren  
 PD Dr. R. Holle, GSF Oberschleißheim

26.08. - 28.08.1999      Bayes-Methoden in der Medizin  
 Dr K. Ickstadt, TU Darmstadt

30.09. - 02.10.1999      Nichtparametrische Verfahren in der Biometrie  
 Prof. Dr. E. Brunner, Universität Göttingen

Weitere Kurse sind in Planung. Die Kurse sind stark anwendungsorientiert und die Lehrinhalte werden mittels praxisnaher Computerübungen vertieft.

Für weitergehende Informationen wenden Sie sich bitte an:

Dr. Katrin Jensen, Abteilung Medizinische Biometrie der Universität Heidelberg

Im Neuenheimer Feld 305, 69120 Heidelberg

Tel: 06221/56-4180, -4141; FAX: 06221/56-4195; E-Mail: jensen@imbi.uni-heidelberg.de

Dr. Birgit Stadler, Zentrum für Studienberatung und Weiterbildung

Friedrich-Ebert-Anlage 22-24, 69117 Heidelberg

Tel: 06221/54-7815, -7810; FAX: 06221/54-7811, E-Mail: Birgit.Stadler@urz.uni-heidelberg.de

<http://www.rzuser.uni-heidelberg.de/~ah5/postgrad.html>

**Ankündigungen und Veranstaltungshinweise**

**Workshop Meta-Analyse vom 05.- 06. November 1999**

**Tutorial Meta-Analyse am 04. November 1999**

**Ort:**

Westfälische Wilhelms-Universität Münster

Beratungsstelle für Organisationen (BfO)

Prof. Dr. H. Holling (Geschäftsführender Direktor)

Fliednerstraße 21

48149 Münster

**Nähere Informationen:** <http://wwwpsy.uni-muenster.de/inst4/AEHolling/metashop.htm>

Themenübersicht

Datengewinnung, Random vs. Fixed Effects-Modelle, Publication Bias,

Effektstärkemaße,

Repeated Measurement-Modelle, Heterogenitätsmodellierung, Prüfung von

Strukturhypothesen,

Bayes-Ansätze, Meta-Analyse in der Qualitätssicherung, Software zur

Meta-Analyse



\*\*\*

**Ankündigung und Call for Papers**

Gemeinsamer Workshop der Arbeitsgruppen

- Statistische Methoden in der Epidemiologie
- Statistische Methoden in der Medizin

der Deutschen Region der Internationalen Biometrischen Gesellschaft und der Arbeitsgruppen

- Statistische Methodik der Klinischen Forschung
- Methoden der Prognose- und Entscheidungsfindung

der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V.  
zum Thema:**Medical Decision Making – methodische Aspekte****18. bis 20. November 1999 in Freiburg**

Deadline für die Anmeldung von Beiträgen ist der **17. September 1999**. Dabei sind ausdrücklich sowohl Vorträge aus theoretisch-mathematischer Sicht als auch aus der praktischen Anwendung erwünscht. Im Rahmen des Workshops wird ein **Tutorium** zum Thema *Multivariable Prozesse: Modelle, Analysen und Interpretationen* stattfinden. Das Tutorium wird von Nanny Wermuth, ZUMA Mannheim, und Reinhold Streit, Psychologisches Institut der Universität Mainz, gehalten.

Beitragsanmeldungen sind an einen Leiter der oben genannten Arbeitsgruppen zu schicken:

Dr. Ulrich Mansmann, FU Berlin, Institut für Medizinische Statistik, Epidemiologie und Informatik, Hindenburgdamm 30, D-12200 Berlin, Tel.: 030/8445-2093, Fax: 030/8445-4471

**Email: mansmann@medizin.fu-berlin.de**

Dr. Willi Sauerbrei, Institut für Medizinische Biometrie und Medizinische Informatik, Klinikum der Universität, Stefan-Meier-Str. 26, D-79104 Freiburg, Tel.: 0761/203-6669, Fax: 0761/203-6680

**Email: wfs@imbi.uni-freiburg.de**

Dr. Karen Steindorf, Knoll AG, MPF/FBK – Biometrie, Postfach 21 08 05, 67008 Ludwigshafen, Tel.: 0621/589-1914, Fax: 0621/589-1925

**Email: karen.steindorf@knoll-ag.de**

Dr. Andreas Ziegler, Philipps-Universität Marburg, Institut für Medizinische Biometrie und Epidemiologie, Bunsenstraße 3, D-35037 Marburg, Tel.: 06421/28-5787, Fax: 06421/28-8921

**Email: ziegler@mail.uni-marburg.de****Organisation vor Ort:**

Dr. Willi Sauerbrei, Institut für Medizinische Biometrie und Medizinische Informatik, Klinikum der Universität, Stefan-Meier-Str. 26, D-79104 Freiburg, Tel.: 0761/203-6669, Fax: 0761/203-6680

**Email: wfs@imbi.uni-freiburg.de**

\*\*\*

**European Course in Advanced Statistics (ECAS)****The seventh course in the ECAS Programme:****ENVIRONMENTAL STATISTICS**

Organised by the Swedish Statistical Association

Sponsored by the European Commission

Location and time: Garpenberg, County of Dalarna, Sweden

Informatik, Biometrie und Epidemiologie in Medizin und Biometrie - Band 30 - Heft 2/1999



September 6 – 10, 1999

ECAS is a programme intended to train graduates and postgraduates in special areas of statistics and to promote new applications of statistics.

**Scientific programme:** A total of approximately 25 lectures will be devoted to the following topics:

**Stochastic modelling vs. Deterministic reductionism in environmental science** (Main speaker: Peter young, UK)

**Sensitivity analysis of environmental models** (Main speaker: Andrea Saltelli, Italy)

**Spatial statistics and applications in environmental science** (Main speaker: Noel Cressie, USA)

**Extreme value theory and applications in environmental science** (Main speaker: Ross Leadbetter, USA)

General information:

Lectures and course materials will be in English.

Registration Fees: Students: 800 SEK

University employees: 2.000 SEK

Other participants: 4.000 SEK

Approved participants must pay the registration fee by June 15, 1999.

Contact Information:

Karin Johansson, Conference Secretary

Department of Mathematics

Linköping University

SE-58183 Linköping, SWEDEN

Tel.: +46 13 281449

Fax: +46 13 100746

Email: [kajoh@mai.liu.se](mailto:kajoh@mai.liu.se) (Karin Johansson, Conference Secretary)

[Angri@mai.liu.se](mailto:Angri@mai.liu.se) (Anders Grimvall, Chairman of the Organising and Scientific Committees)

Web page: <http://www.mai.liu.se/Stat/ecas99.html>

### Hinweis zur Mitgliedsbeitragsrechnung 1999

Sehr geehrte Mitglieder der GMDS,  
liebe Kolleginnen und Kollegen,

vor einigen Tagen sind Ihnen die Mitgliedsbeitragsrechnungen für das Kalenderjahr 1999 zugegangen. In diesem Zusammenhang werden von Ihnen gelegentlich Fragen hinsichtlich der steuerlichen Behandlung von Mitgliedsbeiträgen aufgeworfen. Grundsätzlich können Mitgliedsbeiträge an die GMDS wie Spenden in Ihrer Steuererklärung angegeben werden. Oder auch die Angabe als Werbungskosten ist möglich, wenn auch nur unter bestimmten Voraussetzungen. Wichtig ist, daß die Anrechnung von Mitgliedsbeiträgen durch die Finanzämter nach dem Abflußprinzip erfolgt. Danach ist das Jahr der Zahlung maßgeblich für die Anrechnung in Ihrer Steuererklärung und nicht das Jahr der Rechnungsstellung oder das Kalenderjahr, für das die Rechnung gestellt wurde. So können Sie bspw. auch Mitgliedsbeiträge für das Jahr 1996, die Sie erst in 1999 bezahlen, in Ihrer Steuererklärung für das Jahr 1999 steuermindernd geltend machen. Sie sollten Ihrem Finanzamt die Rechnung oder Zahlungserinnerung zusammen mit einer Kopie des Überweisungsauftrags oder des Kontoauszugs, der die Zahlung bzw. Überweisung oder Abbuchung beim Lastschriftverfahren nachweist, vorlegen. Ich bitte alle Mitglieder, die nicht am Lastschriftverfahren teilnehmen, recht herzlich, die Mitgliedsbeiträge so bald wie möglich zu überweisen. Bitte vergessen Sie auch nicht, Ihre Mitgliedsnummer auf dem Überweisungsträger anzugeben.

Mit freundlichen Grüßen

Ihr

Th. Banasiewicz

(Geschäftsführer und Schatzmeister)



### Neuaufnahmen

Als neue Mitglieder begrüßen wir recht herzlich

Name/Vorname/Akad. Titel	Ort	Name/Vorname/Akad. Titel	Ort
Bachler, Christa, Dr. rer. nat.	Lübeck	Jung, Bettina	Bonn
Berg, Klaudia, Dipl.-Ökotroph.	München	Jürgens, Christine, M.B. Ch. B. (GB)	Münster
Bertling, Frank, Dr. med.	Rheinbach	Langer, Frank, Dr. rer. nat.	Bad Homburg
Boehm, Roland, Dr. med.	München	Löffler, Jürgen, Dr. phil. nat.	Planegg
Eisele, Reiner, Dipl.-Kaufm.	Sinzig	Ness, Günther, Dr. med. Dipl.-Ing.	Darmstadt
Fischer, Imma, Dr. rer. nat..	Tübingen	Ordu, Yusuf, Dr. med.	Bad Waldsee
Franzkwiak de Rodriguez, Martina, Dr. med.	Erlangen	Özdemir, Fatma, Dipl.-Inf. Med.	Rheinfelden
Friedrich, Holger, Dr. med.	Tübingen	Roesner, Arnold	Freiburg
Gerkum, Gamila	Bonn	Schedy, Maria, Dipl.-Med. Inform.	Darmstadt
Giesen, Ludger, Dr. med.	Dinslaken	Seelhorst, Martin, Dipl.-Ing.	Obertshausen
Graf, Johannes	Wachtberg	Snoek, Hendrik, Dipl. Inform. Med.	Lübeck
Gramatzki, Jürgen	Ludwigsburg	Uesbeck, Mechthild, Dipl.-Inform.	Tübingen
Gutenbrunner, Cornelius, Dr. rer. nat.	Marburg	Unkelbach, Oliver	Bornheim-Waldorf
Hempel, Volker, Dr. rer. nat.	Düsseldorf	Verhey, Janko, Dr. rer. nat.	Göttingen
Hoppe, Stephan, Dr. sc. hum.	Neuherberg	Wresky, Hans-Peter, Dr. med.	Michelbach

Ganz herzlich begrüßen wir das nachfolgende Unternehmen als förderndes Mitglied und bedanken uns für die großzügige finanzielle Unterstützung:

Firma	Straße/Postfach	PLZ/Ort
G M T Gesellschaft für Medizintechnik mbH	Ziegelhüttenweg 45	60598 Frankfurt

### Förderpreis für Studierende 1999

## Die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V.

schreibt zur Förderung des wissenschaftlichen Nachwuchses den

### Förderpreis für Studierende 1999

für herausragende Abschlusarbeiten auf den Gebieten Medizinische Informatik, Medizinische Biometrie und Epidemiologie aus.

Der Preis ist mit DM 500,00 dotiert und wird feierlich im Rahmen der 44. Jahrestagung der GMDS 1999 in Heidelberg überreicht. Die Teilnahme an der Tagung ist kostenlos.

Kriterien zur Beurteilung der eingereichten Arbeiten sind sowohl die Relevanz der Ergebnisse für die Medizin als auch die erzielten methodischen Fortschritte auf einem der o.g. Fachgebiete.

Bewerben können sich Absolventen/innen eines Studiengangs einer wissenschaftlichen Hochschule (bspw. Informatik, Mathematik, Medizin) mit Ihrer Abschlusarbeit (z.B. Diplomarbeiten, Medizindissertationen), die 1997, 1998 oder 1999 abgegeben wurde.



Bitte senden Sie zur Bewerbung in jeweils dreifacher Ausfertigung Ihre Arbeit, Ihr Abschlußzeugnis und ein Beurteilungsschreiben des/r Betreuer/in Ihrer Arbeit an die

Geschäftsstelle der GMDS e.V.  
Herbert-Lewin-Str. 1  
D – 50931 Köln

*Einsendeschluß ist der 30. Juni 1999.*

#### Posterpreis 1999

### Die Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V.

schreibt zur Förderung der Wissenschaft den

#### Posterpreis 1999

für herausragende Poster auf den Gebieten Medizinische Informatik, Medizinische Biometrie und Epidemiologie aus.

Das Programmkomitee der GMDS-Jahrestagung 1999 wird auch in diesem Jahr die im Rahmen der 44. Jahrestagung der GMDS in Heidelberg ausgestellten Poster bewerten und einen Posterpreis verleihen. Die Jahrestagung findet vom 13. – 16. September 1999 in Heidelberg statt. Der Posterpreis ist mit DM 500,00 dotiert.

Die Kriterien für die Beurteilung der Poster gliedern sich in zwei Gruppen: Zum einen wird der Inhalt des Posters bewertet. Zum anderen wird gleichrangig die Präsentationsform beurteilt. Hier spielen das graphische Layout, die Konzentration auf das Wesentliche, der geeignete Verweis auf weiterführende Quellen eine wesentliche Rolle. Poster (Plakate) müssen aus bis zu 10 m Entfernung auf sich aufmerksam machen und zum Lesen aus ca. 1 m Entfernung einladen.

#### Briefwahl 1999 zum Präsidium der GMDS

##### Aufruf zur Wahl

Bekanntlich werden seit 1992 die Wahlen zum Präsidium der GMDS im Briefwahlverfahren durchgeführt. In diesem Jahr erfolgt satzungsgemäß die Wahl des ersten Vizepräsidenten, des Schatzmeisters und des Schriftführers, die zuletzt 1997 durchgeführt wurde.

##### Zum Wahlverfahren:

Jedes aktiv wahlberechtigte Mitglied (ordentliches, förderndes und Ehren-Mitglied) der GMDS ist nach § 15 Abs. 1 stimmberechtigt. Studentische Mitglieder sind nicht stimmberechtigt.

Die Unterlagen zur Briefwahl sind diesem Mitteilungsblatt bzw. dieser Ausgabe unserer Zeitschrift beigelegt. Es handelt sich um den Stimmzettel, einen neutralen roten Umschlag und einen weißen Rücksendeumschlag mit der Adresse der GMDS-Geschäftsstelle.

Auf der Seite 1 des Stimmzettels finden Sie eine Anleitung zur ordnungsgemäßen Stimmabgabe, die Sie bitte aufmerksam lesen und befolgen. Die Seite 2 des Stimmzettels enthält eine Tabelle mit den zur Wahl stehenden Ämtern im Präsidium und den Namen der Kandidaten zu den jeweiligen Ämtern.

Mit der Aussendung der beiliegenden Briefwahlunterlagen ist gewährleistet, daß eine Stimmauszählung fristgemäß sechs Wochen vor der kommenden Mitgliederversammlung erfolgt.

Auf der kommenden Mitgliederversammlung findet gemäß § 15 Abs. 5 bekanntlich eine Wahl nur dann noch statt, wenn eine Stichwahl notwendig ist.



Letzter Abgabetermin für die Stimmzettel (Eingang bei der GMDS-Geschäftsstelle): **23. Juli 1999**

Köln, den 26.04.1999

Der Wahlausschuß

**Briefwahl 1999 zum Präsidium der GMDS**  
**Vorstellung der Kandidaten für das Amt des 1. Vizepräsidenten**

**Prof. Dr. Walter Lehmacher**  
 Institut für Medizinische Statistik,  
 Informatik und Epidemiologie  
 der Universität zu Köln  
 Joseph-Stelzmann-Str. 9  
 50931 Köln

### 1. Persönliche und berufliche Angaben

1949	geboren
1968–1973	Studium der Mathematik und Informatik an der Universität Bonn
1973–1977	Wissenschaftlicher Assistent an der Abteilung Statistik der Universität Dortmund; Promotion zum Dr. rer. nat.
1977–1979	Wissenschaftlicher Mitarbeiter im Rechenzentrum der Universität Düsseldorf
1979–1990	Leiter der Arbeitsgruppe Auswertungssoftware im Institut für Medizinische Informatik und Systemforschung (Medis) Gesellschaft für Strahlen- und Umweltforschung (GSF), Neuherberg bei München
1987	Habilitation für das Fachgebiet Biostatistik und Promotion zum Dr. med. habil. an der Medizinischen Fakultät der Universität München
1990–1995	Professor (C4) für Biometrie und Epidemiologie an der Tierärztlichen Hochschule Hannover
seit April 1995	Professor und Direktor des Instituts für Medizinische Statistik, Informatik und Epidemiologie an der Universität zu Köln
	Mitglied des Präsidiums der GMDS (Schriftführer 1989–1991, Sprecher des Fachbereichs Medizinische Biometrie, 1992–1995)

### 2. Ziele für das Amt des ersten Vizepräsidenten der GMDS

Im Interesse aller drei Fachbereiche der GMDS werde ich besonders folgende Ziele verfolgen:

- Ausgewogene Integration aller unserer 3 Fächer bei den Aktivitäten der GMDS, insbesondere bei den Jahrestagungen
- stärkere Einbindung der Arbeitsgruppen in die Jahrestagung
- Repräsentanz der GMDS bei offiziellen Stellen und Mitarbeit bei Gesetzen und Verordnungen im Gesundheitswesen sowie bei der Forschungsförderung
- Kooperation mit anderen wissenschaftlichen Fachgesellschaften, auf nationaler wie auch internationaler Ebene



- Intensivierung der Öffentlichkeitsarbeit und Mitgliederwerbung
- Stärkung der Aktivitäten in der Fort- und Weiterbildung
- Berücksichtigung unserer Fächer bei der Reform des Medizinstudiums sowie in anderen Studiengängen

\*\*\*

**Prof. Dr. Christian Ohmann**

Universität Düsseldorf

Klinik für Allgemeine und Unfallchirurgie

Funktionsbereich Theoretische Chirurgie

Moorenstr. 5

40225 Düsseldorf

**1. Persönliche und berufliche Angaben**

Die Ausbildung umfaßt das Diplom und die Promotion in Mathematik (1976, 1981) sowie das Physikum in Medizin (1984). 1987 erfolgte die Habilitation im Fach Theoretische Chirurgie. Seit 1987 wird der Funktionsbereich Theoretische Chirurgie an der Klinik für Allgemein- und Unfallchirurgie an der Heinrich-Heine-Universität geleitet. Schwerpunkt des Forschungsbereiches sind die Anwendungen medizinischer Entscheidungsfindung in der operativen Medizin sowie kontrollierte klinische Studien in der Chirurgie. Zur Zeit werden medizininformatische Projekte durch die Deutsche Forschungsgemeinschaft, das Bundesministerium für Bildung, Wissenschaft und Forschung, das Bundesgesundheitsministerium und die Europäische Union gefördert. Die Tätigkeiten an den Medizinischen Einrichtungen und der Medizinischen Fakultät der Heinrich-Heine-Universität Düsseldorf umfassen die Leitung der Arbeitsgruppe Klinische Informationsverarbeitung der Medizinischen Einrichtungen, die Beteiligung an der Projektleitung des Krankenhausinformationssystems und die Koordinierung der Einrichtung eines Zentrums für Klinische Studien an der Medizinischen Fakultät.

**2. Ziele für das Amt des ersten Vizepräsidenten der GMDS**

Als theoretisch ausgebildeter Leiter eines praktisch medizinischen Forschungsbereiches verfüge ich über intensive Bezüge zu allen drei Fachbereichen der GMDS. Aus der langjährigen Erfahrung meiner praktischen und wissenschaftlichen Tätigkeit sehe ich folgende Ziele:

- Integration der Fächer Medizinische Informatik, Epidemiologie und Medizinische Biometrie in der praktischen Anwendung (z.B. bei der Implementierung von Koordinierungszentren für Klinische Studien)
- Bessere Verzahnung von Theorie und praktischer medizinischer Anwendung in den drei Fächern durch
  - a) bessere Umsetzung methodischer Ansätze in die praktische Arbeit
  - b) Ableitung von Anstößen aus der praktischen Arbeit für die Theorie.
- Förderung der wissenschaftliche Aspekte von Medizinischer Informatik Medizinischer Biometrie und Epidemiologie durch
  - a) Verbesserung der Drittmittelförderung
  - b) bessere internationale Repräsentanz deutscher Wissenschaftler (Vorträge, Publikationen)
- Förderung der Kooperation mit anderen Fachgesellschaften

\*\*\*



**Prof. Dr. Herbert Witte**

Institut für Medizinische Statistik,  
Informatik und Dokumentation  
Bachstr. 18  
07740 Jena

**1. Persönliche und berufliche Angaben**

1952	geboren
1970–1974	Studium der Technischen Kybernetik (Dipl.-Ing.)
1979	Promotion zum Dr. rer. nat. (Biologie)
1986	Zusatzbezeichnung Fachingenieur für Biomedizinische Technik
1987	Dr. sc. nat. mit einer Arbeit zu methodischen Aspekten der Signalverarbeitung und Modellierung in der neonatologischen Intensivmedizin (Johannes-Müller-Preis der Gesellschaft für Exp. Medizin)
1988/89	Zusatzstudium – Neurologische Abteilung des Thomayer-Krankenhauses in Prag
1989	Facultas docendi für Medizinische Informatik
1990	Wandlung des Dr. sc. nat. in Dr. rer. nat. habil.
1993	Gründungsdirektor und Direktor des Instituts für Medizinische Statistik, Informatik und Dokumentation
1993	Berufung zum C4-Professor
	<u>Gremienarbeit:</u> Sprecher der Klinischen Neurowissenschaften (BMBF) von 1993–1998; Mitglied der Forschungskommission der Fakultät und ab 1999 Vorsitzender; Stellvertr. des Vorsitzenden des Verbundes für Klinische Forschung der Fakultät bis 1998; Mitglied in Senatskommissionen der Universität.
	<u>Fachgesellschaften:</u> Mitglied im ITG-Fachausschuß 6.3 seit 1993; Councilmitglied der EFMI seit 1994; Chairman des IEEE-Chapters EMB (Germany Section) seit 1996; Vorstandsmitglied der DGBMT seit 1997.

**2. Ziele für das Amt des ersten Vizepräsidenten der GMDS**

- Die inhaltliche und organisatorische Zusammenarbeit der in der GMDS vereinigten Fachgebiete muß als Einheit erhalten und weiter vertieft werden.
- Die Stärkung der GMDS als wissenschaftliche Gesellschaft sollte durch den Ausbau der Aktivitäten der Arbeitsgruppen konsequent vorangetrieben werden. Junge Wissenschaftler müssen in den Arbeitsgruppen natürliche Ansprechpartner im Sinne eines Kompetenzzentrums sehen, zu denen unkompliziert eine inhaltliche Bindung aufgebaut werden kann. Es sollten Netzwerke von Arbeitsgruppen aus nationalen und internationalen Fachgesellschaften geschaffen werden, um einen umfassenden Erfahrungsaustausch zu ermöglichen.
- Wesentlich und unverzichtbar ist die stärkere Berücksichtigung der MI als Lehrfach für Medizinstudenten (Approbationsordnung, Studienordnungen der Fakultäten). Dies darf nicht zu Lasten des Lehrdeputats für Biometrie gehen. Ein breiter Erfahrungsaustausch über attraktive Lehrangebote und neue Formen der Lehre in allen Fachgebieten der GMDS sollte unterstützt werden.



**Ankündigung der ISCB-GMDS-Jahrestagung 1999 in Heidelberg**

**ISCB-GMDS-99**

**'Medical Informatics, Clinical Biostatistics and Epidemiology for  
Efficient Health Care and Medical Research'**

13-17 September 1999, Heidelberg, Germany

**20th Annual Conference of the International Society of Clinical Biostatistics (ISCB) 14-17 September**

**joint with the**

**44th Annual Conference of the Deutsche Gesellschaft für Medizinische Informatik, Biometrie  
und Epidemiologie (GMDS) 13-16 September**

held by the University of Heidelberg and the German Cancer Research Center, both located at the Campus „Im Neuenheimer Feld“.

**Scientific program**

The ISCB-GMDS-99 conference is taking on shape. We are planning more than 70 sessions with a total of more than 250 oral and 200 poster presentations covering a broad range of interesting topics in the scientific fields of Medical Informatics, Biostatistics, and Epidemiology. Abstract submission is closed.

The Biostatistics sessions cover a broad range of topics of current interest in design, evaluation and monitoring as well as meta-analysis of clinical trials. Deficits in the implementation of clinical research as well as the impact of guidelines on future drug research will be discussed. The work of Paul Martini, father of clinical research in Germany, will be acknowledged in a separate session. Prognostic models in oncology and frailty models receive increased attention. Amongst other topics, models in preclinical drug development, statistical challenges in molecular medicine and genetics are discussed. We were able to invite a number of internationally renowned speakers.

Topics to be addressed by the Medical Informatics sessions are methods and tools of information processing to support efficient health care and high quality medical research like electronic patient records, medical documentation, co-operative information processing, hospital information systems etc.

The presentations on epidemiology refer to a broad range of recent research topics with emphasis on epidemiology in the elderly (examples include recent studies on dementia or on changes in functional ability), on genetic epidemiology and gene-environment interaction (with application in psychiatric disorders, breast cancer), on epidemiological methods, health systems research, occupational epidemiology and environmental epidemiology (childhood leukemia clustering).

An ISCB-Special-Symposium on "Assessment of Drug Risks" will be held on Friday, 17 September, all day.

*The complete preliminary program will be available by 22 May at <http://www.dkfz-heidelberg.de/biostatistics/iscb-gmds-99>*

**Tutorials**

Informatik, Biometrie und Epidemiologie in Medizin und Biometrie - Band 30 - Heft 2/1999



On **Sunday, 12 September**, three afternoon tutorials will be offered:

Adaptive designs in clinical studies (Peter Bauer, Meinhard Kieser)

Medical controlling (Andreas J. Goldschmidt)

3D visualization in medicine (Andreas Pommert, Ulf Tiede, Karl-Heinz Hühne)

Costs: 225 DM (150 DM for members of GMDS; ISCB, DVMD or BVMI)

On **Monday, 13 September**, three all-day tutorials will be held:

Statistical methods in genomics and computational biology (Richard Simon)

Introduction to frailty models (Philip Hougaard)

Developing and validating clinical prediction models (Frank E. Harrell jr.)

Costs: 450 DM (300 DM for members of GMDS; ISCB, DVMD or BVMI)

### Language

The working language of ISCB-sessions and joint sessions of ISCB and GMDS will be English. The working language of GMDS-sessions is either English or German.

### Registration

Registration fee for the full conference participation covers access to all scientific events and industrial presentations and exhibitions, receipt of conference material including abstract volume, refreshments during breaks and free receptions. Registration can be done preferably by using the online registration form

(<http://www.dkfz-heidelberg.de/biostatistics/iscb-gmds-99>)

or by contacting (also for any further information)

ISCB-GMDS-99 Conference Office

c/o Dept. of Medical Biometry

University of Heidelberg

Im Neuenheimer Feld 305

D-69120 Heidelberg

Germany

FAX +49-6221-56-4195

Phone +49-6221-56-5656

e-mail [ISCB-GMDS-99@dkfz-heidelberg.de](mailto:ISCB-GMDS-99@dkfz-heidelberg.de)

### Registration fees

	Full Conference	ISCB Conference	GMDS Conference
I am member of:			
ISCB and GMDS	DM 380,-	DM 330,-	DM 330,-
ISCB	DM 430,-	DM 330,-	
GMDS	DM 430,-		DM 330,-
neither ISCB nor GMDS	DM 480,-	DM 380,-	DM 380,-
DVMD	DM 200,-*		
Students	DM 50,-*		
1-day ticket	DM 150,-*		

\* GMDS-proceedings not included



In case of registration **before 31 May 1999** the conference fee except for students and 1-Day-tickets is reduced by DM 50,-.

#### **Social program**

We have planned a number of social events that are aimed at bringing together scientists from both societies. The night before the start of the conference, an informal get-together is organized in an old brewery in the old town. There will be receptions in the town hall sponsored by the mayor of Heidelberg and in the Old University Hall sponsored by the rector of the University of Heidelberg, during which an ad-hoc choir that all participants are invited to join will perform Bach. The conference dinner will take place in the Heidelberg castle.

Three different excursions are planned, a boat tour on the Neckar river, a bike trip to Speyer and a bus-tour to the 'Pfalz' with wine-tasting. Guided walks through town have been arranged during the conference. Two 2 day pre/post conference excursions are planned.

#### **Vergünstigter Bezug der Zeitschrift „Technology and Health Care“ für GMDS-Mitglieder**

Der niederländische Verlag IOS Press bietet GMDS-Mitgliedern den vergünstigten Bezug der Zeitschrift „Technology and Health Care“ an. Ein Jahrgang mit 6 Ausgaben im Abonnement kostet für GMDS-Mitglieder NLG 190,00 inklusive Versandkosten. Der normale Preis beträgt zur Zeit NLG 600,00.

Diese Zeitschrift ist das offizielle Organ der European Society for Engineering and Medicine. Sie behandelt die sich überlappenden Themenbereiche der Physik, des Engineerings und der Informatik auf der einen Seite sowie der Humanbiologie, der grundlegenden medizinischen Wissenschaften und klinische Medizin auf der anderen Seite. Ferner stellt sie ein Forum für die Diskussion sozio-ökonomischer Aspekte in Verbindung mit medizinischer Technik bereit.

„Technology and Health Care“ soll ein Forum für die Präsentation neuer Konzepte, Entwicklungen und Verfahren in Verbindung mit der Nutzung von Technik in der Medizin sowie der sozio-ökonomischen Folgen dieser Nutzung sein. Demzufolge beinhaltet die Zeitschrift alle Bereiche des biomedizinischen Engineerings.

Für weitere Informationen wenden Sie sich bitte direkt an  
IOS Press

Van Diemenstraat 94

1013 CN Amsterdam

Netherlands

Tel.: +31 20 638 21 89

Telefax: +31 20 620 34 19

Email: [market@iospress.nl](mailto:market@iospress.nl)

Internet: <http://www.iospress.nl>

## PHARMALOG Institut für klinische Forschung GmbH

Ein unabhängiges europäisches CRO

Das **Pharmalog Institut** als renommiertes **Institut für klinische Studien (CRO)** mit Auftraggebern aus der forschenden Pharmaindustrie sucht für die Leitung der Biometrie-Abteilung

### BIOMETRIKER / DATENMANAGER

Ihre Fähigkeiten sollten sein:

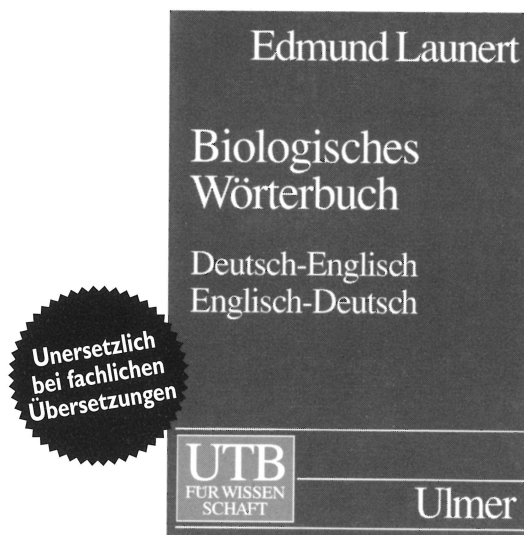
- mehrjährige Erfahrung in statistischer Planung und Auswertung/Reports (ICH)
- Routine in SAS und Englisch
- erfolgsorientierter Führungsstil und straffe Organisation

Unser Team von Projektleitern und Monitoren (ca. 25 Mitarbeiter in München) freut sich auf eine(n) einsatzfreudige(n) Kollegen(in) an der zentralen Schnittstelle „Statistik“.

Bitte schicken Sie Ihre aussagefähigen Bewerbungsunterlagen (mit Lichtbild) an:

**PHARMALOG Institut für klinische Forschung GmbH**, z. Hd. Herrn Eberhardt  
Hermann-Schmid-Straße 10 · D-80336 München · Tel. (0 89) 54 46 37-0 · Fax (0 89) 54 46 37-50  
E-Mail: [Pharmalog@t-online.de](mailto:Pharmalog@t-online.de) · Home Page: <http://www.pharmalog.com>

## Ein umfangreiches Nachschlagewerk.



### Biologisches Wörterbuch.

Deutsch-Englisch /

Englisch-Deutsch.

Edmund Launert. 1998.

739 Seiten.

DM 78,- /

öS 876,- / sFr 71.-.

ISBN 3-8252-8105-1.

Für die Übersetzung von wissenschaftlichen Arbeiten und von biologischer Fachliteratur stellt dieses Wörterbuch **nahezu 45.000 Stichwörter** bereit. Sie betreffen schwerpunktmäßig den Bereich Biologie mit seinen Fachbereichen Zoologie und Botanik mit Pflanzenphysiologie sowie die Genetik und die Verhaltensforschung. Die angrenzenden Gebiete sind in dem Werk ebenfalls berücksichtigt: Medizin, Veterinärmedizin, Bio-

chemie, Ökologie, Bodenkunde, Agrar- und Forstwissenschaften sowie die Grundlagen der Geologie. Besonders zu erwähnen ist die Aufnahme von zahlreichen Trivialnamen von Tieren und Pflanzen, deren Einträge in der Regel zusätzlich durch ihre wissenschaftlichen Namen ergänzt werden. Dieses Nachschlagewerk enthält keine Erklärungen der einzelnen Fachbegriffe, wie sonst im Lexikon üblich, sondern nur ihre Übersetzung.

Bestellen Sie in Ihrer Buchhandlung oder bei:  
Verlag Eugen Ulmer, Postf. 70 05 61, 70574  
Stuttgart, Tel.: 0711 / 4507-121, Fax: 0711 /  
4507-120. Homepage: [www.ulmer.de](http://www.ulmer.de)

### BUCH-COUPON

- ☐ Senden Sie mir „**Dictionary of Nutrition**“ zum Preis von DM 78,- /  
öS 876,- / sFr 71.-. Best.-Nr. 81051.

- ☐ Senden Sie mir kostenlos Ihr  
Gesamtverzeichnis „Ulmenblatt“.

Name/Vorname

Straße/Nr.

PLZ, Ort

Biometrie 2/99

4 schnelle Wege zum Buch: 1. Kauf in  
Ihrer Buchhandlung. 2. Per e-mail: [info@ulmer.de](mailto:info@ulmer.de) 3. Per Telefon (0711) 45 07-  
121. 4. Per Telefax (0711) 45 07-120.



# Ein umfangreiches Nachschlagewerk

*Die englische Sprache ist in der wissenschaftlichen Kommunikation dominant. Ein Großteil des Austausches findet jedoch in der jeweiligen Landessprache statt. Daher ist es sehr wichtig, die gesuchten Fachbegriffe schnell und problemlos in die wichtigsten europäischen Sprachen übersetzen zu können.*

Ca. 3.000  
Fachbegriffe

in fünf  
Sprachen

Claus Leitzmann / Ulrike Dauer

## Dictionary of Nutrition

Wörterbuch der Ernährung

Dictionnaire de la Nutrition

Dizionario di Nutrizione

Diccionario de Nutrición

Ulmer

### Dictionary of Nutrition.

Wörterbuch der  
Ernährung.

Englisch-Französisch-  
Deutsch-Italienisch-  
Spanisch.

Prof. Dr. C. Leitzmann,  
U. Dauer.

2. Auflage 1996.

516 Seiten.

DM 128,-

öS 934,- / sFr 114.-.

ISBN 3-8001-2148-4

Bisher waren Fachbegriffe der Ernährungsphysiologie oft nur unzureichend in einzelne Sprachen übersetzt und schwer auffindbar. Dieses Buch hilft, die fachbezogene Kommunikation auf internationaler Ebene zu erleichtern und zu fördern.

### Zum Buch

Mit diesem Werk liegt ein Verzeichnis von relevanten Begriffen für die Ernährungswissenschaft in fünf wichtigen europäischen Sprachen vor. **DIESES BUCH ENTHÄLT UNGEFÄHR 3000 FACHBEGRIFFE.** Im ersten Teil des Buches werden die englischen Stichwörter in alphabetischer Reihenfolge aufgeführt, in englischer Sprache definiert und in vier Sprachen übersetzt. Der

zweite Teil enthält die Übersetzungen aller Stichwörter, in den vier Sprachen getrennt. Mittels einer Verweisnummer ist der entsprechende Begriff im ersten Teil des Buches zu finden.

### Aus dem Inhalt

3.000 englische Stichwörter mit Definitionen und Übersetzung in 4 Sprachen. Stichwörter in Deutsch, Französisch, Italienisch, Spanisch.

### Der Autor

Prof. Dr. rer. nat. C. Leitzmann arbeitet am Institut für Ernährungswissenschaft der Justus-Liebig-Universität Gießen. Er ist Autor vieler Veröffentlichungen und Mitglied zahlreicher wissenschaftlicher Gesellschaften.

Coupon Ihrer Buchhandlung geben oder senden an: Verlag Eugen Ulmer, Postfach 70 05 61, 70574 Stuttgart. Fax: 0711/4507-120

### BUCH-COUPON

☐ Senden Sie mir das Buch „**Dictionary of Nutrition**“ zum Preis von DM 128,- / öS 934,- / sFr 114.-. Best.-Nr. 21484.

☐ Senden Sie mir kostenlos Ihr Gesamtverzeichnis „Ulmenblatt“

Datum/Unterschrift

Name/Vorname

Straße/Nr.

PLZ, Ort

Biometrie 2/99



VERLAG  
EUGEN  
ULMER